



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 94257

TO: Phillip Gambel
Location: CM1/9E12/8B03
Art Unit: 1644
Wednesday, May 28, 2003

Case Serial Number: 919408

From: Susan Hanley
Location: Biotech-Chem Library
CM1-6B05
Phone: 305-4053

susan.hanley@uspto.gov

Search Notes

94257

From: STIC-ILL
Sent: Monday, May 19, 2003 8:32 AM
To: STIC-Biotech/ChemLib
Subject: FW: 09/919408 lemiskchka amd

-----Original Message-----

From: Gambel, Phillip
Sent: Monday, May 19, 2003 8:29 AM
T : STIC-ILL
Subject: 09/919408 lemiskchka amd

stic

please perform a sequence and a sequence interference search for

SEQ ID NO:1

SEQ ID NO: 2

SEQ ID NO:3

SEQ ID NO:4

hard copy

thanx

phillip gambel
art unit 1644
308-3997

1644 mailbox 9e12

CREF

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 25, 2003, 02:59:08 ; Search time 8656.83 seconds
(without alignments)
11608.404 Million cell updates/sec

Title: US-09-919-408-1
Perfect score: 3453
Sequence: 1 GCGCCCTGGCTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_nam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3453	100.0	3453	6	AR005211	AR005211 Sequence
2	3453	100.0	3453	6	AR071703	AR071703 Sequence
3	3453	100.0	3453	6	I25169	I25169 Sequence 1
4	3453	100.0	3453	6	I40600	I40600 Sequence 1
5	3426	99.2	3426	10	MUSFLK2	M64689 Mouse flk-2
6	3344	96.8	3521	6	I44732	I44732 Sequence 1
7	3344	96.8	3521	10	MMFLI3	X59398 Mouse Flt3
8	2227.2	64.5	3501	6	AR005212	AR005212 Sequence
9	2227.2	64.5	3501	6	AR071704	AR071704 Sequence
10	2227.2	64.5	3501	6	I25170	I25170 Sequence 3
11	2227.2	64.5	3501	6	I40601	I40601 Sequence 3
12	2204.8	63.9	3476	6	I40109	I40109 Sequence 1
13	2203	63.8	3475	6	I44733	I44733 Sequence 3
14	2203	63.8	3475	9	HS002687	U02687 Human growth
15	2191.4	63.5	3422	9	HSFLT3RTK	Z26652 H.sapiens F
16	2081.8	60.3	3120	6	I44518	I44518 Sequence 22
17	1206	34.9	1894	6	AR149571	AR149571 Sequence
18	474.6	13.7	181003	2	AC127349	AC127349 Mus muscu
19	427.4	12.4	429	10	MUS3RTK	L36163 Mus musculu
20	363.2	10.5	480	9	HUM3RTK	L36162 Homo sapien
21	357.4	10.4	3992	6	AX394288	AX394288 Sequence
22	357.4	10.4	3992	9	HSCFMS	X03663 Human mRNA
23	349.8	10.1	3545	5	AF153446	AF153446 Danio rer
24	344.2	10.0	389	10	AY094358	AY094358 Rattus no
25	341.8	9.9	4577	14	FCSSMONC	K01643 Feline sarc
26	340.2	9.9	3828	4	CATEMSC	J03149 Cat (F.dome
27	336.8	9.8	4222	4	S76596	S76596 c-kit=recep
28	335.6	9.7	3705	10	RRCKTIRTK	X62491 R.rattus mr
29	332.6	9.6	2176	4	AF263826	AF263826 Bos tauru
30	332.6	9.6	2176	4	AF263827	AF263827 Bos tauru
31	332.6	9.6	3069	4	BOVCKR	D16680 Bovine mRNA
32	331.4	9.6	2919	6	A82459	A82459 Sequence 47
33	331.4	9.6	2960	4	SSAJ3228	AJ223228 Sus scrof
34	331.4	9.6	2960	4	SSAJ3229	AJ223229 Sus scrof
35	331.4	9.6	2960	4	SSAJ3230	AJ223230 Sus scrof
36	331.2	9.6	3816	10	RATCKITPO	D12524 Rattus norv
37	330.4	9.6	3679	10	RRCSE1	X61479 Rat mRNA fo
38	330.2	9.6	2894	4	ECA224643	AJ224643 Equus cab
39	330.2	9.6	2894	4	ECA224644	AJ224644 Equus cab
40	330.2	9.6	2973	4	ECA224642	AJ224642 Equus cab
41	330	9.6	3154	4	AF044249	AF044249 Canis fam
42	329.8	9.6	2979	4	AF448148	AF448148 Canis fam
43	329.2	9.5	2894	4	ECA224645	AJ224645 Equus cab
44	329	9.5	5084	6	AX195908	AX195908 Sequence
45	329	9.5	5084	6	AX331941	AX331941 Sequence

ALIGNMENTS

RESULT 1
AR005211
LOCUS AR005211 3453 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5747651.
ACCESSION AR005211
VERSION AR005211.1 GI:3966090
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka,I.R.
TITLE Antibodies against tyrosine kinase receptor flk-1
JOURNAL Patent: US 5747651-A 1 05-MAY-1998;
FEATURES Location/Qualifiers

Db 2041 CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100
Qy 2101 ATTTTGAATATTGTTGCTATGTTGCTGACCTCCTCACTACCTAAGAAGTAAAGAGAGAAG 2160
Db 2101 ATTTTGAATATTGTTGCTATGTTGCTGACCTCCTCACTACCTAAGAAGTAAAGAGAGAAG 2160
Qy 2161 TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAAATTTCACTTCTTACCTTACTTTC 2220
Db 2161 TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAAATTTCACTTCTTACCTTACTTTC 2220
Qy 2221 CAGGCACATTCAATTTCCAGCATGCTGTTTACAGGAAGTTCAGTTTACACCCGCCCTTG 2280
Db 2221 CAGGCACATTCAATTTCCAGCATGCTGTTTACAGGAAGTTCAGTTTACACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCACTTCTGAAAGTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCACTTCTGAAAGTGAATATGAA 2340
Qy 2341 AACCAGAAGAGCTGGCAGAGAAGAGAGAGATTTGAACGTCGAGTTTGAAGAC 2400
Db 2341 AACCAGAAGAGCTGGCAGAGAAGAGAGAGATTTGAACGTCGAGTTTGAAGAC 2400
Qy 2401 CTCCTTTGCTTTGCTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAAGTCGTGT 2460
Db 2401 CTCCTTTGCTTTGCTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAAGTCGTGT 2460
Qy 2461 GTCACAGAGACCTGGCAGCCAGGAATGTTGGTACACCACGGGAAGTGGTGAAGATC 2520
Db 2461 GTCACAGAGACCTGGCAGCCAGGAATGTTGGTACACCACGGGAAGTGGTGAAGATC 2520
Qy 2521 TGTGACTTTGGACTGGCCCGACAGATCCTCAGCGACTCCAGCTACGTCGAGGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGACAGATCCTCAGCGACTCCAGCTACGTCGAGGGGCAAC 2580
Qy 2581 GCACGGCTGCGGTGAAGTGGATGGCCAGCAGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2581 GCACGGCTGCGGTGAAGTGGATGGCCAGCAGAGCTTATTTGAAGGATCTACACAATC 2640
Qy 2641 AAGAGTGAGCTGTGGTCTCAGCGATCCTTCTCGGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTGAGCTGTGGTCTCAGCGATCCTTCTCGGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CCTTACCTGGCATCTCTGTCGAGCTAATCTTATATAACTGATTCAGAGTGGATTAAA 2760
Db 2701 CCTTACCTGGCATCTCTGTCGAGCTAATCTTATATAACTGATTCAGAGTGGATTAAA 2760
Qy 2761 ATGGAGCAGCATCTATGCCAGAGAGGATATCTTTGTATGCAATCCTGCTGGGCT 2820
Db 2761 ATGGAGCAGCATCTATGCCAGAGAGGATATCTTTGTATGCAATCCTGCTGGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCATCCTTCCCAACTGACTTCAATTTTAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCCTTCCCAACTGACTTCAATTTTAGGATGTCAGCTG 2880
Qy 2881 GCAGAGCAGAGAGCATGTATCAGAACATCCTATCCATCTACCAAAACAGGGCGCCCT 2940
Db 2881 GCAGAGCAGAGAGCATGTATCAGAACATCCTATCCATCTACCAAAACAGGGCGCCCT 2940
Qy 2941 CAGCAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGTTTACAGAGAA 3000
Db 2941 CAGCAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGTTTACAGAGAA 3000
Qy 3001 AGAAGTTAGCGAGGAGGCTTTGGACCCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGGAGGCTTTGGACCCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA 3060
Qy 3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCTGCTGGACTTTTCTCT 3120
Db 3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCTGCTGGACTTTTCTCT 3120
Qy 3121 AGATGCTGTCTGCCATTTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180

Db 3121 AGATGCTGTCTGCCATTTACTCCAAAGTGACTTCTATAAATAAAACCTCTCTCTCGCACAG 3180
Qy 3181 GCGGGAGAGCCAAATAATGAGACTTGTGGTGAGCCCGCTACCCCTGGGGCCCTTTCCACG 3240
Db 3181 GCGGGAGAGCCAAATAATGAGACTTGTGGTGAGCCCGCTACCCCTGGGGCCCTTTCCACG 3240
Qy 3241 AGCTTGAGGGGAAAGCCATGTATCTGAATAATAGTATATCTTGTAAATAACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAAGCCATGTATCTGAATAATAGTATATCTTGTAAATAACGTGAACAA 3300
Qy 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTTAA 3360
Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTTAA 3360
Qy 3361 AATACTATGTAACCTTTTTCATCTATTTAGTATATATTTATGATGGAATAAACTTTC 3420
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTATATATTTATGATGGAATAAACTTTC 3420
Qy 3421 TACTGTAAAAAATAAAAAAATAAAAAA 3453
Db 3421 TACTGTAAAAAATAAAAAAATAAAAAA 3453

RESULT 2
AR071703
LOCUS AR071703 3453 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5912133.
ACCESSION AR071703
VERSION AR071703.1 GI:7222591
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka,I.R.
TITLE Method for isolating stem cells expressing flk-1 receptors
JOURNAL Patent: US 5912133-A 1 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..3453 /organism="unknown"
BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN

Query Match 100.0%; Score 3453; DB 6; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCTGGCTACCGCGCTCCGGAGGCGCATCGGGGCTTGGCGCAGCGCAGCGACCGG 60
Db 1 GCGGCTGGCTACCGCGCTCCGGAGGCGCATCGGGGCTTGGCGCAGCGCAGCGACCGG 60
Qy 61 GCGGCTGGCTGCTGTTGTTTGTCAATAATGATTTTGAGACCGTTTACAAACCAAGAC 120
Db 61 GCGGCTGGCTGCTGTTGTTTGTCAATAATGATTTTGAGACCGTTTACAAACCAAGAC 120
Qy 121 CTGCTGTGATCAAGTGTGTTTTATCAGTCATGAGAACATGCTCATCAGCGGAAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTTATCAGTCATGAGAACATGCTCATCAGCGGAAAG 180
Qy 181 CCATCATCGTACCGAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Db 181 CCATCATCGTACCGAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Qy 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300
Db 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300
Qy 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGACCTTTTCTGCTCTGCGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGACCTTTTCTGCTCTGCGGTCTTTAAGCAC 360
Qy 361 AGCTCCCTGGGCTGCCAGCCGCTTGTATTTACAAACAGAGGATCTCTTCCATGGCC 420


```
Qy 2581 GCACGGCTGCCGTTGAAGTGGATGCGACCCGAGAGCTTATTGGAAGGATCTACACAATC 2640
Db 2581 GCACGGCTGCCGTTGAAGTGGATGCGACCCGAGAGCTTATTGGAAGGATCTACACAATC 2640
Qy 2641 AAGAGTGACGCTGGTCTACGGCATCCTCTCTGGGAGATATTTCACCTGGGTGTAAC 2700
Db 2641 AAGAGTGACGCTGGTCTACGGCATCCTCTCTGGGAGATATTTCACCTGGGTGTAAC 2700
Qy 2701 CCTTACCTGGCATTCCTGTGCGAGCTAACTTCTATATAACTGATTTCAGAGTGGATTAAA 2760
Db 2701 CCTTACCTGGCATTCCTGTGCGAGCTAACTTCTATATAACTGATTTCAGAGTGGATTAAA 2760
Qy 2761 ATGGAGCAGCATTCTATGCCACAGAGGGATATCTTTGTAATGCAATCCTGTGGCT 2820
Db 2761 ATGGAGCAGCATTCTATGCCACAGAGGGATATCTTTGTAATGCAATCCTGTGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCCATCCTTCCCAACCTGACTTCATTTTATGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCCATCCTTCCCAACCTGACTTCATTTTATGATGTCAGCTG 2880
Qy 2881 GCAGAGCAGAAAGAGTATGATCAGAAATCCATCCATCTACCAAAACAGCGGCCCT 2940
Db 2881 GCAGAGCAGAAAGAGTATGATCAGAAATCCATCCATCTACCAAAACAGCGGCCCT 2940
Qy 2941 CAGCAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTCAACAGAA 3000
Db 2941 CAGCAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTCAACAGAA 3000
Qy 3001 AGAAGTTAGCAGAGCGGCTTGGACCCCGCCACCTACAGGCTGTAGACCCAGAGCA 3060
Db 3001 AGAAGTTAGCAGAGCGGCTTGGACCCCGCCACCTACAGGCTGTAGACCCAGAGCA 3060
Qy 3061 AGATTAGCTCGCTCTGAGGAAGCGCCTACAGCGGCTGTGCTCGCTGGACTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAAGCGCCTACAGCGGCTGTGCTCGCTGGACTTTCTCT 3120
Qy 3121 AGATGCTGTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCACAG 3180
Db 3121 AGATGCTGTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCACAG 3180
Qy 3181 CGGGAGAGCCAAATAGAGACTTGTGGTGAGCCCGCTACCCCTGGGGGCTTTCCAG 3240
Db 3181 CGGGAGAGCCAAATAGAGACTTGTGGTGAGCCCGCTACCCCTGGGGGCTTTCCAG 3240
Qy 3241 AGCTTGAGGGGAAAGCCATGTCTGAAATAGTATATCTGTAATACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAAGCCATGTCTGAAATAGTATATCTGTAATACGTGAACAA 3300
Qy 3301 ACCAAACCCGTTTTTGTAAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360
Db 3301 ACCAAACCCGTTTTTGTAAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360
Qy 3361 AATACTATGTAACCTTTTCACTATTTAGTATATATTTATGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAACCTTTTCACTATTTAGTATATATTTATGATGGAATAAATCTTC 3420
Qy 3421 TACTGTAAAAAAGGAAAAAAAAAAAAAAAAAAAA 3453
Db 3421 TACTGTAAAAAAGGAAAAAAAAAAAAAAAAAAAA 3453
```

```
RESULT 3
LOCUS 125169 3453 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 1 from patent US 5548065.
ACCESSION 125169
VERSION 125169.1 GI:1605039
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka, I.R.
```

```
TITLE Tyrosine kinase receptor human flk-2-specific antibodies
JOURNAL Patent: US 5548065-A 1 20-AUG-1996;
FEATURES
Location/Qualifiers
Source 1..3453
BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN
Query Match 100.0%; Score 3453; DB 6; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGGCTGGCTACCGCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGACGCGCGG 60
Db 1 GCGGCTGGCTACCGCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGACGCGCGG 60
Qy 61 CGGCTGCTGCTGCTTTGTTTTCAGTAAGTATCTTGAGACCCCTTACAAACCAAGAC 120
Db 61 CGGCTGCTGCTGCTTTGTTTTCAGTAAGTATCTTGAGACCCCTTACAAACCAAGAC 120
Qy 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180
Db 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180
Qy 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Db 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Qy 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTGGTGGTCCATC 300
Db 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTGGTGGTCCATC 300
Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360
Qy 361 AGCTCCCTGGGTGCGAGCGCCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 361 AGCTCCCTGGGTGCGAGCGCCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Qy 421 ATCTTGAAGCTCAGAGACCCAGGAGGAGATACTACTCTCATATTTCAGAGCGAACGC 480
Db 421 ATCTTGAAGCTCAGAGACCCAGGAGGAGATACTACTCTCATATTTCAGAGCGAACGC 480
Qy 481 GCCAACTACACAGTACTGTTTCAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAACTACACAGTACTGTTTCAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540
Qy 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCATCTGCTCATCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCATCTGCTCATCTCCGAGGGTGT 600
Qy 601 CCGGAGCCCATCTGTGGAGTGGGTCTCGAGCTCCACAGGAAAGCTGTAAGAGAA 660
Db 601 CCGGAGCCCATCTGTGGAGTGGGTCTCGAGCTCCACAGGAAAGCTGTAAGAGAA 660
Qy 661 GGCCTGCTGTTGTCAGAAAGGAGGAAGGTTACTTCATGAGTGTTCGGAACAGACATC 720
Db 661 GGCCTGCTGTTGTCAGAAAGGAGGAAGGTTACTTCATGAGTGTTCGGAACAGACATC 720
Qy 721 AGATGCTGTGCTAGAAATGCATCTGGCCCGCAATGCACCAAGCTGTTCCATAGATCTA 780
Db 721 AGATGCTGTGCTAGAAATGCATCTGGCCCGCAATGCACCAAGCTGTTCCATAGATCTA 780
Qy 781 AACCAGGCTCCTCAGAGCACACTGCCAGTATTCTCTGAAAGTGGGGAAACCTTTG 840
Db 781 AACCAGGCTCCTCAGAGCACACTGCCAGTATTCTCTGAAAGTGGGGAAACCTTTG 840
Qy 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTTCGGGCTACCTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTTCGGGCTACCTGGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGGAGGAGGCGAGCTACTTTTGTGATGAGTACCTTACTCCAAACAGGACCATG 960
```

|||||
901 AAAGCCCTGGAGGCGAGCTACTTTGAGATGAGTACTACTCACAAACAGGACCATG 960
QY
961 ATTGGATTCTTTGGCCCTTTGTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
Db
961 ATTGGATTCTTTGGCCCTTTGTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
QY
1021 TGCTCTTCTCAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080
Db
1021 TGCTCTTCTCAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080
QY
1081 ATAAACGCTACCAAGCTCGCAAGAGAGATGAATTTGACCCCGTAGAAAAAGTTCTGCTTC 1140
Db
1081 ATAAACGCTACCAAGCTCGCAAGAGAGATGAATTTGACCCCGTAGAAAAAGTTCTGCTTC 1140
QY
1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGACGTGGATCTTCTCAAGCCTCA 1200
Db
1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGACGTGGATCTTCTCAAGCCTCA 1200
QY
1201 TTTCTCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Db
1201 TTTCTCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY
1261 AAGAACAGCCAGGAGGTACATATTTATGCAAGAAATGATGACGCCAGTTCAACAAA 1320
Db
1261 AAGAACAGCCAGGAGGTACATATTTATGCAAGAAATGATGACGCCAGTTCAACAAA 1320
QY
1321 ATGTTCAAGCTGAATATGAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Db
1321 ATGTTCAAGCTGAATATGAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
QY
1381 CGCTCTCTGCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTTGTCGGAC 1440
Db
1381 CGCTCTCTGCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTTGTCGGAC 1440
QY
1441 AAATCTCCCAATTCACAGGAGAAATCCAGAAAGAGTTTGGAAATAAAGAGCTTAACAGA 1500
Db
1441 AAATCTCCCAATTCACAGGAGAAATCCAGAAAGAGTTTGGAAATAAAGAGCTTAACAGA 1500
QY
1501 AAAGTGTGGCCAGTGGTTCGAGCAGTACTCTAAATATCAGTGAGCCCGGGAAGGG 1560
Db
1501 AAAGTGTGGCCAGTGGTTCGAGCAGTACTCTAAATATCAGTGAGCCCGGGAAGGG 1560
QY
1561 CTCTGGTCAATGTGTCGTACAAATCTATGGGCAGCTCTTGCGAAACCATCTTTTAA 1620
Db
1561 CTCTGGTCAATGTGTCGTACAAATCTATGGGCAGCTCTTGCGAAACCATCTTTTAA 1620
QY
1621 AACTCACCAGGCCCTTCCCTTTTCAATCAAGAACACATCTCTCTATGCGACCATTTGG 1680
Db
1621 AACTCACCAGGCCCTTCCCTTTTCAATCAAGAACACATCTCTCTATGCGACCATTTGG 1680
QY
1681 CTCTGCTCCCTTCATTTGTTCTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740
Db
1681 CTCTGCTCCCTTCATTTGTTCTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740
QY
1741 TTTAGGTACGAGACGTCGAGATGATCAGGTGACTGGCCCTCGGATAACGAGTAC 1800
Db
1741 TTTAGGTACGAGACGTCGAGATGATCAGGTGACTGGCCCTCGGATAACGAGTAC 1800
QY
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860
Db
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860
QY
1861 TTAGAGTTTGGGAAGGCTCTGGGGTCTGGCGCTTTTCGGGAGGTTGATGAAACGCCAGGCC 1920
Db
1861 TTAGAGTTTGGGAAGGCTCTGGGGTCTGGCGCTTTTCGGGAGGTTGATGAAACGCCAGGCC 1920
QY
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTCAAGATGCTAAAGAGAAA 1980
Db
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTCAAGATGCTAAAGAGAAA 1980
QY
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
|||||

Db
1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGATGACCCACCTGGGA 2040
QY
2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTACAGGGCCAGTGTACTTG 2100
Db
2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTACAGGGCCAGTGTACTTG 2100
QY
2101 ATTTTGTGAATATTTGTTGCTATGTTGACCTCTCACTACTAAGAAAGTAAAAAGAGAGAAG 2160
Db
2101 ATTTTGTGAATATTTGTTGCTATGTTGACCTCTCACTACTAAGAAAGTAAAAAGAGAGAAG 2160
QY
2161 TTTACAGACATGACAGAGATTTTTAAGAACATAATTTTTCAGTTCTTACCCACTTTTC 2220
Db
2161 TTTACAGACATGACAGAGATTTTTAAGAACATAATTTTTCAGTTCTTACCCACTTTTC 2220
QY
2221 CAGGCACATTTCAATTTCCAGCATGCTGTTTCCAGGAGAGTTTCACTTACCCCGCCCTTG 2280
Db
2221 CAGGCACATTTCAATTTCCAGCATGCTGTTTCCAGGAGAGTTTCACTTACCCCGCCCTTG 2280
QY
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
Db
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
QY
2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTGAAGAC 2400
Db
2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTGAAGAC 2400
QY
2401 CTCCTTTTCTGCTTACCAAGTGGCCAAAGGCATGGAATTTCTGGAGTTCAAGTCTG 2460
Db
2401 CTCCTTTTCTGCTTACCAAGTGGCCAAAGGCATGGAATTTCTGGAGTTCAAGTCTG 2460
QY
2461 GTCCACAGAGACTGTGCGCAGGAGATGTGTTGGTCACCCAGGGAAGTGTGGAAGATC 2520
Db
2461 GTCCACAGAGACTGTGCGCAGGAGATGTGTTGGTCACCCAGGGAAGTGTGGAAGATC 2520
QY
2521 TGTGACTTTGGACTGGCCGAGACATCTCTGAGCGACTCCAGCTACGCTGTCAGGGGCAAC 2580
Db
2521 TGTGACTTTGGACTGGCCGAGACATCTCTGAGCGACTCCAGCTACGCTGTCAGGGGCAAC 2580
QY
2581 GCACGCTGCGCGTGAAGTGGATGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
Db
2581 GCACGCTGCGCGTGAAGTGGATGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
QY
2641 AAGAGTGAGCTGTGCTTACGGCATCTCTCTGCGGAGATATTTTCACTGGGTGTGAAC 2700
Db
2641 AAGAGTGAGCTGTGCTTACGGCATCTCTCTGCGGAGATATTTTCACTGGGTGTGAAC 2700
QY
2701 CCTTACCCTGGCATCTCTGTCAGCGTAACTTCTATAAAGTATGATCAGAGTGAATTTAA 2760
Db
2701 CCTTACCCTGGCATCTCTGTCAGCGTAACTTCTATAAAGTATGATCAGAGTGAATTTAA 2760
QY
2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTTGATGCAATCTCTGCTGGCT 2820
Db
2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTTGATGCAATCTCTGCTGGCT 2820
QY
2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGSATGTACGCTG 2880
Db
2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGSATGTACGCTG 2880
QY
2881 GCAGAGGCAAGAGAGATGATCAGAACATCCATCTACCAAAACAGGGCGGCCCT 2940
Db
2881 GCAGAGGCAAGAGAGATGATCAGAACATCCATCTACCAAAACAGGGCGGCCCT 2940
QY
2941 CAGCAGAGAGCGGCTCAGAGCCCGCTCAGAGCCCGCTAGAGAGGATTTACACAGAA 3000
Db
2941 CAGCAGAGAGCGGCTCAGAGCCCGCTCAGAGCCCGCTAGAGAGGATTTACACAGAA 3000
QY
3001 AGAAGTTACGAGGAGCGCTTGGACCCCGCCCTAGCAGGCTGTAGACCCGAGAGCCA 3060
Db
3001 AGAAGTTACGAGGAGCGCTTGGACCCCGCCCTAGCAGGCTGTAGACCCGAGAGCCA 3060
QY
3061 AGATTAGCCTCTGAGGAGCGCCCTACAGCGCTGTGCTTCGCTGGACTTTTCTCT 3120
Db
3061 AGATTAGCCTCTGAGGAGCGCCCTACAGCGCTGTGCTTCGCTGGACTTTTCTCT 3120

Qy 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATATAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATATAATCAAACTCTCTCGCACAG 3180
Qy 3181 GCGGAGAGGCAATAATAGAGACTTGTGGTGAGCCCGCCTACCCCTGGGGGCTTTCCACG 3240
Db 3181 GCGGAGAGGCAATAATAGAGACTTGTGGTGAGCCCGCCTACCCCTGGGGGCTTTCCACG 3240
Qy 3241 AGCTTGAGGGGAAGCCATGTAATAGTAAATAGTATATCTTCTTAATAGCTGAACAA 3300
Db 3241 AGCTTGAGGGGAAGCCATGTAATAGTAAATAGTATATCTTCTTAATAGCTGAACAA 3300
Qy 3301 ACCAAACCGCTTTTCTGAAGGAAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360
Db 3301 ACCAAACCGCTTTTCTGAAGGAAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360
Qy 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATATTTATGATGAAATAAATTC 3420
Db 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATATTTATGATGAAATAAATTC 3420
Qy 3421 TACTGTAATAAAAAAAAAAAAAAAAAAAAAA 3453
Db 3421 TACTGTAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 4
I40600 I40600 3453 bp DNA linear PAT 13-MAY-1997
LOCUS I40600
DEFINITION Sequence 1 from patent US 5621090.
ACCESSION I40600
VERSION I40600.1 GI:2082892
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka, I. R.
TITLE Nucleic acids encoding soluble human FLK-2 extracellular domain
JOURNAL Patent: US 5621090-A 1 15-APR-1997;
FEATURES
Location/Qualifiers
1..3453
/organism="unknown"
BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN

Query Match 100.0%; Score 3453; DB 6; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCTGGCTACCGCGCTCCGGAGGCCATGGGGGGTGGCGAGCGGAGCGGCGG 60
Db 1 GCGGCTGGCTACCGCGCTCCGGAGGCCATGGGGGGTGGCGAGCGGAGCGGCGG 60
Qy 61 CGGCTGCTGCTGCTGTTGTTTCTCAGTAAATGATCTTGAGACCGTTTACAAACCAAGAC 120
Db 61 CGGCTGCTGCTGCTGTTGTTTCTCAGTAAATGATCTTGAGACCGTTTACAAACCAAGAC 120
Qy 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG 180
Db 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG 180
Qy 181 CCATCATGTAACCGAATGGTGGAGAGTCCCGAGAACCTCCAGTGTACCCGAGGGCG 240
Db 181 CCATCATGTAACCGAATGGTGGAGAGTCCCGAGAACCTCCAGTGTACCCGAGGGCG 240
Qy 241 CAGAGTGAAGGAGCGGTATGAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGAGCGGTATGAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300
Qy 301 ACCGTGAAGTGACGTCGCCACCCAGGGGACCTTTCTCGCTCTGCTGGTCTTTAAGCAC 360
Db 301 ACCGTGAAGTGACGTCGCCACCCAGGGGACCTTTCTCGCTCTGCTGGTCTTTAAGCAC 360

Qy 361 AGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 361 AGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Qy 421 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCTATATTCAGAGCGAACGC 480
Db 421 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCTATATTCAGAGCGAACGC 480
Qy 481 GCCAACTACAGTACTGTTCCACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAACTACAGTACTGTTCCACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
Qy 541 AGACCTTACTTTAGGAAGATGAAAACAGAGTGCCTGCTGATCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAAGATGAAAACAGAGTGCCTGCTGATCTCCGAGGGTGT 600
Qy 601 CCGGAGCCACTGTGGAGTGGGTGCTGTCAGCTCCCACAGGAAAGCTGTAAAGAA 660
Db 601 CCGGAGCCACTGTGGAGTGGGTGCTGTCAGCTCCCACAGGAAAGCTGTAAAGAA 660
Qy 661 GGGCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCTATGAGTTGTTCCGGAACAGACATC 720
Db 661 GGGCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCTATGAGTTGTTCCGGAACAGACATC 720
Qy 721 AGATGCTGCTAGAAATGCACCTGGCCCGCAATGCACCAAGCTGTTCCACCATAGATCTA 780
Db 721 AGATGCTGCTAGAAATGCACCTGGCCCGCAATGCACCAAGCTGTTCCACCATAGATCTA 780
Qy 781 AACCAGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840
Db 781 AACCAGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840
Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGGAGGAGGCGACTTCTGAGATGAGTACCTACTCCACAACAGGACCATG 960
Db 901 AAAGCCCTGGAGGAGGCGACTTCTGAGATGAGTACCTACTCCACAACAGGACCATG 960
Qy 961 ATTCGGATTCCTTGGCCCTTTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
Db 961 ATTCGGATTCCTTGGCCCTTTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
Qy 1021 TGCTCTTCTCAAAAGCACCCAGCGTGGTGAGCCATCTCTAGAAAAGGGTTT 1080
Db 1021 TGCTCTTCTCAAAAGCACCCAGCGTGGTGAGCCATCTCTAGAAAAGGGTTT 1080
Qy 1081 ATAAAGCTTACAGCTCCAGAGAGATGAAATGACCCCGTACGAAAGTTCTGCTTC 1140
Db 1081 ATAAAGCTTACAGCTCCAGAGAGATGAAATGACCCCGTACGAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGTGGATCTTCTCAAGCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGTGGATCTTCTCAAGCCTCA 1200
Qy 1201 TTTCTTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
Db 1201 TTTCTTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
Qy 1261 AAGAACAGCCAGGAGAGTACATATCTATGCAGAAATGATGACGCCCGCTACCAAA 1320
Db 1261 AAGAACAGCCAGGAGAGTACATATCTATGCAGAAATGATGACGCCCGCTACCAAA 1320
Qy 1321 ATGTTACAGCTCAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAG 1380
Db 1321 ATGTTACAGCTCAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAG 1380
Qy 1381 GCGTCTCTCTGATGGCTACCCGCTACCCCTCTTGAGACCTGGAAGAAGTTTTCGGAC 1440
Db 1381 GCGTCTCTCTGATGGCTACCCGCTACCCCTCTTGAGACCTGGAAGAAGTTTTCGGAC 1440

1441 AAATCTCCAAATTCACGAGGAAATCCAGAAAGGAGTTGGAAATAAAAAGGCTAACAGA 1500
1441 AAATCTCCAAATTCACGAGGAAATCCAGAAAGGAGTTGGAAATAAAAAGGCTAACAGA 1500
1501 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTAAATATAGATGAGGCGGGAAGGG 1560
1501 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTAAATATAGATGAGGCGGGAAGGG 1560
1561 CTCTCGTCAAAATGCTGCGTACAAATCTATGGGCAGCTCTTGGCAAAACCATCTTTTAA 1620
1561 CTCTCGTCAAAATGCTGCGTACAAATCTATGGGCAGCTCTTGGCAAAACCATCTTTTAA 1620
1621 AACTCACCAGGCCCCCTCCCTTTTCATCCAAAGAACACATCTCTTATGCGACCATTTGGG 1680
1621 AACTCACCAGGCCCCCTCCCTTTTCATCCAAAGAACACATCTCTTATGCGACCATTTGGG 1680
1681 CTCTCTCCCTTCATTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1740
1681 CTCTCTCCCTTCATTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1740
1741 TTTAGGTACGAGTACGATGATCCAGGTGACTGGGCCCTGGATACCGATAC 1800
1741 TTTAGGTACGAGTACGATGATCCAGGTGACTGGGCCCTGGATACCGATAC 1800
1801 TTCTACGTTGACTTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860
1801 TTCTACGTTGACTTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860
1861 TTAGAGTTGGGAAGTCTGGGGTCTGGCGTCTTTCGGGAGGTTGATGAACCCAGCGCC 1920
1861 TTAGAGTTGGGAAGTCTGGGGTCTGGCGTCTTTCGGGAGGTTGATGAACCCAGCGCC 1920
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGTCAAGATGCTAAAGAGAAA 1980
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGTCAAGATGCTAAAGAGAAA 1980
1981 GCTGACAGTGTGAAAAAGAGTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGTGTGAAAAAGAGTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACATCGTGAATCTGTCGGGGATGCACACTGTACAGGCGCATGTACTTTG 2100
2041 CACCATGACATCGTGAATCTGTCGGGGATGCACACTGTACAGGCGCATGTACTTTG 2100
2101 ATTTTGTGAATTTGTTGATGTTGTCACCTCTCACTACCTAAGAGTAAAGAGAGAG 2160
2101 ATTTTGTGAATTTGTTGATGTTGTCACCTCTCACTACCTAAGAGTAAAGAGAGAG 2160
2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220
2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220
2221 CAGGCACATTCAAATTCAGCATGCTGGTTTACAGAGAGTTTCAAGTTACACCCGCCCTTG 2280
2221 CAGGCACATTCAAATTCAGCATGCTGGTTTACAGAGAGTTTCAAGTTACACCCGCCCTTG 2280
2281 GATGAGTCTCAGGTTTCAATGGGAATTCATTTCAATTCATTCGAGATGAGATGAATGAA 2340
2281 GATGAGTCTCAGGTTTCAATGGGAATTCATTTCAATTCATTCGAGATGAGATGAATGAA 2340
2341 AACACAGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2341 AACACAGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2401 CTCTCTTTCGTTTCGATACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCGTGT 2460
2401 CTCTCTTTCGTTTCGATACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCGTGT 2460
2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTACCCACGGAAGGTGTGAGATGC 2520
2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTACCCACGGAAGGTGTGAGATGC 2520
2521 TGTGACTTTGGACTGGCCGAGACATCTCTGACGACTCCAGCTACGTCCTCAGGGGCAAC 2580

2521 TGTGACTTTGGACTGGCCGAGACATCTCTGAGGACTCCAGCTACGTCGTCAGGGGCAAC 2580
2581 GCACGGCTCCGGTGAAGTGGATGGACCCGAGAGCTTATTGAAAGGATCTACACAATC 2640
2581 GCACGGCTCCGGTGAAGTGGATGGACCCGAGAGCTTATTGAAAGGATCTACACAATC 2640
2641 AAGAGTGAGCTGTGGTCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
2641 AAGAGTGAGCTGTGGTCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
2701 CCTTACCCCTGGCATTCCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAA 2760
2701 CCTTACCCCTGGCATTCCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAA 2760
2761 ATGGAGCAGCATTCATGCCACAGAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
2761 ATGGAGCAGCATTCATGCCACAGAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
2821 TTTGACTCAAGGAAGCGGCATCTCTCCCACTGACTTTCATTTTAGGATGTCAGCTG 2880
2821 TTTGACTCAAGGAAGCGGCATCTCTCCCACTGACTTTCATTTTAGGATGTCAGCTG 2880
2881 GCAGAGGCAAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
2881 GCAGAGGCAAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACAGCGCCAGGTGAGATTCACAGAA 3000
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACAGCGCCAGGTGAGATTCACAGAA 3000
3001 AGAAGTTACCGAGGAGCGCTTGGACCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 3060
3001 AGAAGTTACCGAGGAGCGCTTGGACCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 3060
3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCGCGTGTCTCGCTGGACTTTCTCT 3120
3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCGCGTGTCTCGCTGGACTTTCTCT 3120
3121 AGATGCTGTCTGCCATTACTCCAAAGTACTTCTATAAAATCAAACTCTCCTCGCACAG 3180
3121 AGATGCTGTCTGCCATTACTCCAAAGTACTTCTATAAAATCAAACTCTCCTCGCACAG 3180
3181 GCGGGAGAGCCAAATATGAGACTTGTGTGAGCCGCCCTACCCCTGGGGGCTTTCCACG 3240
3181 GCGGGAGAGCCAAATATGAGACTTGTGTGAGCCGCCCTACCCCTGGGGGCTTTCCACG 3240
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300
3301 ACCAAACCCGTTTTTTGCTAAGGGAAGCTAAATATGATTTTTTAAATATGTTTTTAA 3360
3301 ACCAAACCCGTTTTTTGCTAAGGGAAGCTAAATATGATTTTTTAAATATGTTTTTAA 3360
3361 AATACTATCTAATTTTTTCTATCTATTTAGTGAATATTTTATGATGGAATAAATCTTC 3420
3361 AATACTATCTAATTTTTTCTATCTATTTAGTGAATATTTTATGATGGAATAAATCTTC 3420
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 5

MUSFLK2
LOCUS Mouse flk-2 mRNA, complete cds.
DEFINITION
ACCESSION M64689
VERSION M64689.1 GI:193327
KEYWORDS
SOURCE Mus musculus (strain C3H/He)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3426)
Matthews, W., Jordan, C.T., Wiegand, G.W., Pardoll, D. and
Lemischka, I.R.

A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations
Cell 65 (7), 1143-1152 (1991)

Location/Qualifiers

1. 3426

```
1. 1.0420
/organism="Mus
```

$$\text{strain} = \frac{\text{C3H}}{\text{He}}$$

```
/db_xref="tax
```

1. .3426

```
/gene="flk-2"
```

31. 3009

```
/gene="flk-2"
```

```

/codon_start=1
/protein_id="AAA37634.1"
/db_ref="GI:193328"
/translation="MKRLAQRSDRLRLLLLVLSVMILEVTVNQDLPVKICVLISHENN
GSSAKPSSYRMVRSGPEDLQCTPRQSGTVYEAAIVAESGSTLQVLQATPGDL
SCLVTHKHSGLQFDLQNRIGSMALINTVETQAGVLLIHSERANYVLFTVFN
VDTOLVLRREYFRKMNDDALGICSEGVPTVEWGLCSHSRECKEGPAVVRKE
EVLHGLFTGTDICARNALGRECTKLTITDLNQAQPSLTLPOLFLAVGPELWIRCKAI
KYNHGEFLTWEELDKALRGSEYENASTYETNTEMIJLLAFYSSVGRNQTGYTCSSS
KPPSGALVTILEKGITNATISQEEETIDEYKFKSVSRKATPIRTWTFISQASQ
CBQSDGYSISKEDCNKPKGEYIFYAENDDAQTKMTINIRKQPVLANASAQ
ACSGEDGYLPSTWTKCKSDPNCETPEEYGGWNNKANKRVFGWVSSTLMSBAG
KGLLVKCAAYNSMGTSCETIFLNSGPPFPFIDQNTIFYTIGLCLPFIVLVLICHK
YKQRFESOLOMIOVTGPDYFNAYEYFVDFRDEYDLKWFEPRENLEBFKVLGSGAFGR
VNNAATGYISKTCVSIQAVKMLKKKADSCKEALMSKLMMTHLGHNDINVLNLCAG
TILSGVYPLIFECVCGDILNLYLSRKEKPHRTWTEIFKEHNPSYPTTFQAHNSMPG
SREVLPHLDQLDSFGNGMSITSEDEIETENOKRLAEEDDEENLVITLEDLFCATQV
AKMGFELPFGSKVHRDLAARNVLTHGKVLGQIDFGDLADSDSYVYVGRGNARLVPK
WNAPSEFLGFIYTIKSDWMSYGIILLWEIFSLGVNYPYGPIDVANTYKLLIQSGFKMBQF
FYATGIEYVMQSCWAFDSKRRKPSFNLN^SFLLCQCLAEAEACIRTSIHLPKQAAPQ
RGLRAQSPOROVKTHRSR"

```

920 a	822 c	875 a	809 t

Query Match

Query Match : qq 28: score 3426: DB 10: Length 3426.

Query Match	55.20	Score	3420
Best Local Similarity	100.08	Pred No	0

[illegible]

QY I GGGCCCTAGCCGCGCTCCGGAAGGCCATGCGGGCGTTGGCCAGCCAGCCGACCGG GG

61

62

QY 01 CGGC1GCG1GCT1GTTGTTT1GTCAGTAAATGAT1CTTGAGAGCCGT1ACAAACCAAGAC 120

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918

QY 121 CTGGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAATGGCTCATCAGCGGGAAG 180

181 CCATCATCGTACCGAATGGTGGAGGATCCCCAGAAGACCTCCAGTGTACCCCGAGGCGC 240
QY

QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300

QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGACCTTTCCTGCCCTCTGGGCTCTTAAGCAC 360

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTACAAACAGAGGAATCGTTTCCATGGCC 420

361	AGCTCCCTGGGCTGCCACCGCCACTTTTGATTATCAAAAACAGAGGAATCGTTTCCATCGCC	420
421	ATCTTGAACGTGACAGAGACCCAGGACGAGAGAAATACCTACTCCATATTCAGAGCGAACGC	480
421	ATCTTGAACGTGACAGAGACCCAGGACGAGAGAAATACCTACTCCATATTCAGAGCGAACGC	480
481	GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTGCTAAGG	540
481	GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTGCTAAGG	540
541	AGACCTTACTTTAGGAAGATGGAAAAACGAGTGCACCTGCTGTGCATCTCCGAGGGGTGT	600
541	AGACCTTACTTTAGGAAGATGGAAAAACGAGTGCACCTGCTGTGCATCTCCGAGGGGTGT	600
601	CCGAGGCCACTGTGGAGTGGGTCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA	660
601	CCGAGGCCACTGTGGAGTGGGTCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA	660
661	GGCCCTGCTGTTCAGAAAAGGAGAAAGTACTTTCATGAGTTGTTTCGGAACAGACATC	720
721	AGATGCTGCTAGAAATGCATGGGCCGCGAATGCACCAAGCTGTTCCACATAGATCTA	780
721	AGATGCTGCTAGAAATGCATGGGCCGCGAATGCACCAAGCTGTTCCACATAGATCTA	780
781	AACGAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTGAAAGTGGGGGAACCCCTGTGG	840
781	AACGAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTGAAAGTGGGGGAACCCCTGTGG	840
841	ATCAGGTGTAAAGCCATCCATGTGAACCATTGGATTTCGGGCTCAGCTGGGAGCTGGAAG	900
841	ATCAGGTGTAAAGCCATCCATGTGAACCATTGGATTTCGGGCTCAGCTGGGAGCTGGAAG	900
901	AAAGCCCTGGAGGGGACGCTACTTTGAGTAGTAGTACTCCACAACAGGACCATG	960
901	AAAGCCCTGGAGGGGACGCTACTTTGAGTAGTAGTACTCCACAACAGGACCATG	960
961	ATTCGGATTCTTTGGCCTTTGTCTTCGGTGGGAAGAACGACCGGATATTACACC	1020
961	ATTCGGATTCTTTGGCCTTTGTCTTCGGTGGGAAGAACGACCGGATATTACACC	1020
1021	TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT	1080
1021	TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT	1080
1081	ATAAACGCTACCAAGCTCGCAAGAGAGTATGAATTCACCCGTCAGAAAAAGTTCTCGTTC	1140
1081	ATAAACGCTACCAAGCTCGCAAGAGAGTATGAATTCACCCGTCAGAAAAAGTTCTCGTTC	1140
1141	TCAGTCAGGTTTAAAGCGTACCCACGAAATCCGATGCACGTGGATCTCTCTCAAGCCTCA	1200
1141	TCAGTCAGGTTTAAAGCGTACCCACGAAATCCGATGCACGTGGATCTCTCTCAAGCCTCA	1200
1201	TTTCTTGTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
1201	TTTCTTGTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
1261	AAGAACAGCCAGGAGAGTACATATTCTATTCAGAAAAATGATAGCCGCCAGTTCCACCAA	1320
1261	AAGAACAGCCAGGAGAGTACATATTCTATTCAGAAAAATGATAGCCGCCAGTTCCACCAA	1320
1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380
1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380
1381	CGGTCCTGTTCTCTGTATGGCTACCCGCTACCCCTCTTTGGACCTGGAAAGAGTGTTCGGAC	1440
1381	CGGTCCTGTTCTCTGTATGGCTACCCGCTACCCCTCTTTGGACCTGGAAAGAGTGTTCGGAC	1440
1441	AAATCTCCCAATTCACGAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA	1500
1441	AAATCTCCCAATTCACGAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA	1500

```
QY 1501 AAAGTGTGGCCAGTGGGTGCGAGCAGTACTCTAATATACAGTGAGCGCGGGAAGGG 1560
Db 1501 AAAGTGTGGCCAGTGGGTGCGAGCAGTACTCTAATATACAGTGAGCGCGGGAAGGG 1560
QY 1561 CTTCTGGTCAAAATGCTGTGCGGTACAAATCTATGGCAGCTCTTGCGAAACCATCTTTTAA 1620
Db 1561 CTTCTGGTCAAAATGCTGTGCGGTACAAATCTATGGCAGCTCTTGCGAAACCATCTTTTAA 1620
QY 1621 AACTCACAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTTATGCGACCAATGGG 1680
Db 1621 AACTCACAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTTATGCGACCAATGGG 1680
QY 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
QY 1741 TTTAGGTACGAGTACAGTGCAGATGATCAGGTGACTGGCCCCCTGGATAACAGGTAC 1800
Db 1741 TTTAGGTACGAGTACAGTGCAGATGATCAGGTGACTGGCCCCCTGGATAACAGGTAC 1800
QY 1801 TTCTACGTTGACTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCGCCGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCGCCGAGAGAAC 1860
QY 1861 TTAGAGTTTGGAAAGTCTGGGCTGCGCTTTCGGAGGGTGTGAACGCCACGGCC 1920
Db 1861 TTAGAGTTTGGAAAGTCTGGGCTGCGCTTTCGGAGGGTGTGAACGCCACGGCC 1920
QY 1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACACATCGTAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGTACTTG 2100
Db 2041 CACCATGACACATCGTAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGTACTTG 2100
QY 2101 ATTTTGAATATTTGCTATGCTGACCTCTCAACTACCTAAGAAATGAAAGAGAGAG 2160
Db 2101 ATTTTGAATATTTGCTATGCTGACCTCTCAACTACCTAAGAAATGAAAGAGAGAG 2160
QY 2161 TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTCAGTCTTACCCCTACTTTC 2220
Db 2161 TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTACGTTCTTACCCCTACTTTC 2220
QY 2221 CAGGCACATTCAAATTCACGATGCTGGTTCACGAGAGTTCAGTTACACCCGCCCTTG 2280
Db 2221 CAGGCACATTCAAATTCACGATGCTGGTTCACGAGAGTTCAGTTACACCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGTTCAATGGGAATCAATTCATCTGAAGATGAGATGCAATATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTCATCTGAAGATGAGATGCAATATGAA 2340
QY 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400
Db 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400
QY 2401 CTCTCTTCTGCTTCCCTACCAAGTGGCCAAAGCATGGATTCCTGGAGTTCAGTCTGCT 2460
Db 2401 CTCTCTTCTGCTTCCCTACCAAGTGGCCAAAGCATGGATTCCTGGAGTTCAGTCTGCT 2460
QY 2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCACCCAGGAAGGTGGTGAAGATC 2520
Db 2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCACCCAGGAAGGTGGTGAAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2580
```

```
QY 2581 GCACGGCTGCCGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGATCTACAAATC 2640
Db 2581 GCACGGCTGCCGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGATCTACAAATC 2640
QY 2641 AAGAGTACGCTGTGTCCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGAAC 2700
Db 2641 AAGAGTACGCTGTGTCCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGAAC 2700
QY 2701 CTTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAACTGATTCAGAGTGAATAAA 2760
Db 2701 CTTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAACTGATTCAGAGTGAATAAA 2760
QY 2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
Db 2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCATTTTTAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCATTTTTAGGATGTCAGCTG 2880
QY 2881 GCAGAGGAGAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGGAGAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGATTCACAGAGAA 3000
Db 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGATTCACAGAGAA 3000
QY 3001 AGAAGTTAGCGAGGAGGCTTGACCCCGCCACCTAGCAGGTGTAGACCCGAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGGAGGCTTGACCCCGCCACCTAGCAGGTGTAGACCCGAGAGCCA 3060
QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCGTGTGCTTCGCTGAGCTTTCTCT 3120
Db 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCGTGTGCTTCGCTGAGCTTTCTCT 3120
QY 3121 AGATGCTGTCTGCATTTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTGTCTGCATTTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
QY 3181 GCGGGAGAGCCAAATATGAGACTTTGTTGGTGAGCCCGCTACCTGGGGCCCTTTCCACG 3240
Db 3181 GCGGGAGAGCCAAATATGAGACTTTGTTGGTGAGCCCGCTACCTGGGGCCCTTTCCACG 3240
QY 3241 ACCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300
Db 3241 ACCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCGCTTTTGTGCTAAGGGAAGCTAAATATGATTTTAAATCTATGTTTAA 3360
Db 3301 ACCAAACCGCTTTTGTGCTAAGGGAAGCTAAATATGATTTTAAATCTATGTTTAA 3360
QY 3361 AATACTATGTAATCTTTTCTATCTATTTAGTGATATATTTATGATGGAATAAATTTTC 3420
Db 3361 AATACTATGTAATCTTTTCTATCTATTTAGTGATATATTTATGATGGAATAAATTTTC 3420
QY 3421 TACTGT 3426
Db 3421 TACTGT 3426
```

RESULT 6

```
LOCUS I44732
DEFINITION Sequence 1 from patent US 5635388.
ACCESSION I44732
VERSION I44732.1 GI:2469445
KEYWORDS
SOURCE Unknown.
ORGANISM Unknow.
REFERENCE 1 (bases 1 to 3521)
AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.
```


Db 2032 GCTGACAGCTGTGAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2091
Qy 2041 CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTCTAGGCGCCAGTGTACTTG 2100
Db 2092 CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTCTAGGCGCCAGTGTACTTG 2151
Qy 2101 ATTTTGAATATTTGCTATGCTGATGGTACCTCTCAACTACCTAAGAAAGTAAAGAGAGAG 2160
Db 2152 ATTTTGAATATTTGCTATGCTGATGGTACCTCTCAACTACCTAAGAAAGTAAAGAGAGAG 2211
Qy 2161 TTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCAGTCTTACCTACTTTC 2220
Db 2212 TTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCAGTCTTACCTACTTTC 2271
Qy 2221 CAGGACATTTCAATTTCCAGATGCTGTGTTTCACGAGAGTTTCAGTTACACCCGCCCTTG 2280
Db 2272 CAGGACATTTCAATTTCCAGATGCTGTGTTTCACGAGAGTTTCAGTTACACCCGCCCTTG 2331
Qy 2281 GATCAGCTCTCAGGTTTCAATTTGGAATTTCAATTTCTGAAGATGAGATTGAATATGAA 2340
Db 2332 GATCAGCTCTCAGGTTTCAATTTGGAATTTCAATTTCTGAAGATGAGATTGAATATGAA 2391
Qy 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAAGCTGCTGACGTTTGAAGAC 2400
Db 2392 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAAGCTGCTGACGTTTGAAGAC 2451
Qy 2401 CTCCTTTGCTTTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCGTGT 2460
Db 2452 CTCCTTTGCTTTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCGTGT 2511
Qy 2461 GTCCACAGAGACCTGGCCAGCAGGAATGTGTTGTCACCCAGCGGAAGTGGTGAAGATC 2520
Db 2512 GTCCACAGAGACCTGGCCAGCAGGAATGTGTTGTCACCCAGCGGAAGTGGTGAAGATC 2571
Qy 2521 TGTGACTTTGACTGGCCCGAGACATCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC 2580
Db 2572 TGTGACTTTGACTGGCCCGAGACATCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC 2631
Qy 2581 GCAGGCTGCGGTCAGTGGATGCGCCAGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2632 GCAGGCTGCGGTCAGTGGATGCGCCAGAGCTTATTTGAAGGATCTACACAATC 2691
Qy 2641 AAGAGTACGCTGCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2692 AAGAGTACGCTGCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2751
Qy 2701 CCTTACCTGGCATCTGCTGACGCTAACCTTCTATAAAGTATTCAGAGTGATTTAA 2760
Db 2752 CCTTACCTGGCATCTGCTGACGCTAACCTTCTATAAAGTATTCAGAGTGATTTAA 2811
Qy 2761 ATGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCCTGCTGGCT 2820
Db 2812 ATGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCCTGCTGGCT 2871
Qy 2821 TTTGACTCAAGGAGCGGCTCTCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880
Db 2872 TTTGACTCAAGGAGCGGCTCTCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2931
Qy 2881 GCAGAGCAGAGAGC-----ATGTATCAGAACATCCATCC 2917
Db 2932 GCAGAGCAGAGAGAGCGATGTATCAGAACATGGTGGCAGCGTCCCAAGAACATCCATCC 2991
Qy 2918 ATCTACCAAAACAGCGCGCTCTACGAGAGGCGGCTCAGAGCC-CAGTCCGACAG 2976
Db 2992 ATCTACCAAAACAGCGCGCTCTACGAGAGGCGGCTCAGAGCGGCTTGGACCGCCACAG 3051
Qy 2977 CGCCAGGTGAAGATTCACAGAGAAAGATGAGCAGGAGGCTTTGGACCGCCGCCCT 3036
Db 3052 GCCCAGGTGAAGATTCACAGAGAAAGATGAGCAGGAGGCTTTGGACCGCCGCCCT 3111
Qy 3037 AGCAGGCTGTAGACCGCAGAGCCAGATTTAGCCTCGCTCTGAGGAAGCGCCCTACAGG 3096
Db 3112 AGCAGGCTGTAGACCGCAGAGCCAGATTTAGCCTCGCTCTGAGGAAGCGCCCTACAGG 3171

Qy 3097 CGTTGCTTGGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3156
Db 3172 CGTTGCTTGGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3231
Qy 3157 AAAATCAAAACCTCTCTCGACAGGGGGAGAGCCAAATATGAGACTTTGTTGTGAGCCC 3216
Db 3232 AAAATCAAAACCTCTCTCGACAGGGGGAGAGCCAAATATGAGACTTTGTTGTGAGCCC 3291
Qy 3217 GCCTACCTGGGGCCCTTT-----CCACAGCTTGGAGGAAAGCCATGATCTGAA 3268
Db 3292 GCCTACCTGGGGCCCTTTCCAGGCCGCCAGGCTTGGAGGAAAGCCATGATCTGAA 3351
Qy 3269 ATATAGTATATTTCTGTAATACGTGAACAAACCAACCCCGTTTTTTGTAAGGAAAG 3328
Db 3352 ATATAGTATATTTCTGTAATACGTGAACAAACCAACCCCGTTTTTTGTAAGGAAAG 3411
Qy 3329 CTAATATGATTTTTTAAATACTATGTTTTTAAATACTATGTAACCTTTTTCATCTATTTA 3388
Db 3412 CTAATATGATTTTTTAAATACTATGTTTTTAAATACTATGTAACCTTTTTCATCTATTTA 3471
Qy 3389 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAA 3438
Db 3472 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAA 3521

RESULT 7

MMFLT3
LOCUS Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF. 3521 bp mRNA linear ROD 04-DEC-1991
DEFINITION X59398
ACCESSION X59398.1 GI:50978
VERSION
KEYWORDS Flt3 gene; tyrosine kinase receptor.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1321)
AUTHORS Rosnet O.
TITLE Direct Submission.
JOURNAL Submitted (07-MAY-1991) O. Rosnet, Cancerologie et therapeutique exp., Inst National de la Sante, 27 Boulevard Lei Roure, 13009 Marseille, France
REFERENCE 2 (bases 1 to 3521)
AUTHORS Rosnet O., Marchetto S., delapeyriere O. and Birnbaum D.
TITLE Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family
JOURNAL Oncogene 6 (9), 1641-1650 (1991)
MEDLINE 92019834
PUBMED 1656368.

FEATURES

source

1..3521

/organism="Mus musculus"

/db_xref="taxon:10090"

1..3521

/gene="Flt3"

1..3521

/gene="Flt3"

/evidence-experimental

82..3084

/gene="Flt3"

/codon_start=1

/product="Flt3 protein"

/protein_id="CAA42041.1"

/db_xref="GI:50979"

/db_xref="MGD:MGI:95559"

/db_xref="SWISS-PROT:Q00342"

/translation="MRALQSRDRRLLLVLSVMILETNTNODLPVICKVLISHENN
GSSAGKPSYRMVRCSPEDLOCTPRQSEGTVEATVEAESGSLQLVQLATPGDL
SLCWFKHSSLCQPHFDLQNRGIVSMAILNVTETAGELVLLHIQSEARNYVLFVN
VDTQLYLVRRPYFRKMENODALLCISIEGPVTEWVLCSSHRSKEGPAVRKE
EKLHLEFQDIRCCARNALGRESTKLFITDLNQAPQSLPOLFLKVGPEPLTRCKAI
HVNHGFGLTLEDKALEBGSVFENSTYTNRTMIRILLAFVSSVGRNDTGYTTCSSS

QY	1921	TATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAAAGAGAAA	1980			
DB	1972	TATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAAAGAGAAA	2031			
QY	1981	GCTGACAGCTGTGAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA	2040			
DB	2032	GCTGACAGCTGTGAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA	2091			
QY	2041	CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTCAGGCCAGTGTACTTG	2100			
DB	2092	CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTCAGGCCAGTGTACTTG	2151			
QY	2101	ATTTTCAATATTTGTTGCTATGTTGACCTCTCTCAACTACCTAAGAAGTAAAGAGAGAAG	2160			
DB	2152	ATTTTCAATATTTGTTGCTATGTTGACCTCTCTCAACTACCTAAGAAGTAAAGAGAGAAG	2211			
QY	2161	TTTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCASTTCTTACCCCTACTTC	2220			
DB	2212	TTTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCASTTCTTACCCCTACTTC	2271			
QY	2221	CAGGCATTCAAATTCACGATCCCTGGTTCACGAGAGTTCAGTTACACCCGCCCTTG	2280			
DB	2272	CAGGCATTCAAATTCACGATCCCTGGTTCACGAGAGTTCAGTTACACCCGCCCTTG	2331			
QY	2281	GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTCAAGATGAGATTGAATATGAA	2340			
DB	2332	GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTCAAGATGAGATTGAATATGAA	2391			
QY	2341	AACCAGAGAGGCTGGCAGAGAGAGAGAGATTTGAACGTCGACGTTTGAAGAC	2400			
DB	2392	AACCAGAGAGGCTGGCAGAGAGAGAGAGATTTGAACGTCGACGTTTGAAGAC	2451			
QY	2401	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTG	2460			
DB	2452	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTG	2511			
QY	2461	GTCCACAGAGACCTGGCAGCAGGAATGTGTGTCACCCAGGAGGTTGGTGAAGATC	2520			
DB	2512	GTCCACAGAGACCTGGCAGCAGGAATGTGTGTCACCCAGGAGGTTGGTGAAGATC	2571			
QY	2521	TGTGACTTTTGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC	2580			
DB	2572	TGTGACTTTTGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC	2631			
QY	2581	GCACGGCTGCGGTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACAATC	2640			
DB	2632	GCACGGCTGCGGTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACAATC	2691			
QY	2641	AAGAGTGACGCTGGTCTACGGCATCCTTCTCAGGAGATATTTTCACTGGGTGTGAAC	2700			
DB	2692	AAGAGTGACGCTGGTCTACGGCATCCTTCTCAGGAGATATTTTCACTGGGTGTGAAC	2751			
QY	2701	CCTTACCCTGGCATTCCTGTGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA	2760			
DB	2752	CCTTACCCTGGCATTCCTGTGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA	2811			
QY	2761	ATGAGAGACCATTCATGTCACAGAGAGGATATACCTTTGTAATGCAATCCTGCTGGGCT	2820			
DB	2812	ATGAGAGACCATTCATGTCACAGAGAGGATATACCTTTGTAATGCAATCCTGCTGGGCT	2871			
QY	2821	TTTTCACATCAAGAGAGCGGCATCTTCCCAACCTGACTTCAATTTTATAGGATGTGAGTG	2880			
DB	2872	TTTTCACATCAAGAGAGCGGCATCTTCCCAACCTGACTTCAATTTTATAGGATGTGAGTG	2931			
QY	2881	GCAGAGGAGAGAGAGC-----ATGATCAGAAACATCCATCC	2917			
DB	2932	GCAGAGGAGAGAGAGC-----ATGATCAGAAACATCCATCC	2991			
QY	2918	ATCTACCAAAACAGCGGCCCTCTCAGCAGAGAGCGGGCTCAGAGCC-CAGTCCGCCACAG	2976			
DB	2992	ATCTACCAAAACAGCGGCCCTCTCAGCAGAGAGCGGGCTCAGAGCCCGCCATCGCCACAG	3051			
QY	2977	CGCCAGGTGAAGATTCACAGAGAAAGATTTAGCGAGGAGGCTTGGACCCCGCCACCT	3036			
DB	3052	CGCCAGGTGAAGATTCACAGAGAAAGATTTAGCGAGGAGGCTTGGACCCCGCCACCT	3111			
QY	3037	AGCAGGCTGTAGACCGCAGAGCCAGATAGCTCGCCTCTCAGGAAGCCCTACAGCG	3096			
DB	3112	AGCAGGCTGTAGACCGCAGAGCCAGATAGCTCGCCTCTCAGGAAGCCCTACAGCG	3171			
QY	3097	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTCTCTGCCATTACTCCAAAGTACTTCTAT	3156			
DB	3172	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTCTCTGCCATTACTCCAAAGTACTTCTAT	3231			
QY	3157	AAATCAAACTCTCTCCACAGCGGAGAGCCAAATATGAGACTTGTGGTGAGCCC	3216			
DB	3232	AAATCAAACTCTCTCCACAGCGGAGAGCCAAATATGAGACTTGTGGTGAGCCC	3291			
QY	3217	GCCTACCCCTGGGGCTTT-----CCACGAGCTTGAGGGAAAGCCATGTATCTGAA	3268			
DB	3292	GCCTACCCCTGGGGCTTT-----CCACGAGCTTGAGGGAAAGCCATGTATCTGAA	3351			
QY	3269	ATATAGTATATCTTGTAAATACGTGAACAAACCAACCCCGTTTTTGTCTAAGGAAAG	3328			
DB	3352	ATATAGTATATCTTGTAAATACGTGAACAAACCAACCCCGTTTTTGTCTAAGGAAAG	3411			
QY	3329	CTAAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTTCATCTATTTA	3388			
DB	3412	CTAAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTTCATCTATTTA	3471			
QY	3389	GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAAA	3438			
DB	3472	GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAAA	3521			
RESULT 8						
AR005212						
LOCUS	AR005212	Sequence 3501 bp	DNA linear PAT 04-DEC-1998			
DEFINITION	Sequence 3 from patent US 5747651.					
ACCESSION	AR005212					
VERSION	AR005212.1 GI:3966091					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3501)					
AUTHORS	Lemischka, I.R.					
TITLE	Antibodies against tyrosine kinase receptor flk-1					
JOURNAL	Patent: US 5747651-A 3 05-MAY-1998;					
FEATURES	Location/Qualifiers					
source	1..3501					
BASE COUNT	1068 a	709 c	784 g 940 t			
ORIGIN	/organism="unknown"					
Query Match 64.5%; Score 2227.2; DB 6; Length 3501;						
Best Local Similarity 79.6%; Pred. No. 0;						
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;						
QY	8	GGCTACCGCGCTCGGAGGCCATCGCGCGTGGCGAGCGCAGCGAGCCGCCGCGGCTGC	67			
DB	35	GGGGACCCCGGCTCCGGAGGCCATCGCGCTGGC---GCAGCAGCGGGCGCCGCTGC	91			
QY	68	TGCTGCTGTTGTTTCTCAGTAATGATTTCTTGAGACCGTTTACAAACCAAGACCTGCCTG	127			
DB	92	CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTTACAAATCAAGATCTGCGCTG	151			
QY	128	TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAAGCCATCAT	187			
DB	152	TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGATTCATCAGTGGGAAAGTCAATCAT	211			
QY	188	CGTACCAATGTTGCGGAGGATCCCAAGAGACCTCCAGTGTACCCCGAGCGGAGAGTG	247			
DB	212	CATATCCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTGCGTTTGAGACCCCGAGAGCT	271			

QY 248 AAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGGAGTCTGGTCCATCACCTGC 307
DB 272 CAGGGACAGTGTACGAAGCTGCCGTGTGAAGTGGATGTATCTGCTTCCATCACACTGC 331
QY 308 AAGTGCAGCTGCCACCCAGGAGGACCTTTCTCCCTCTGGTCTTTTAAAGCAGAGTCC 367
DB 332 AAGTGTGTGATGCGCCAGGGAACATTTCTGTCTCTGGTCTTTAAGCAGAGTCC 391
QY 368 TGGCTGCCAGCGCACTTTGATTTACAAAAGAGAAATCGTTCCATGCCATCTTGA 427
DB 392 TGAATTTGCCAGCCACATTTTGTATTTACAAAAGAGAGTGTGTTCATGGTCAATTTGA 451
QY 428 ACCTGACAGACCCAGGAGGAGATACCTACTCCATATTCAGAGGGAACCGGCCAACT 487
DB 452 AATGACAGAAACCAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT 511
QY 488 ACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGGAGACCTT 547
DB 512 ACACAATATTTTACAGTGAATTAAGAAATACCCCTGCTTTACACAATTAAGAAGACCTT 571
QY 548 ACTTTAGGAAGATGGAACACGAGTGCATGCTCTGCATCTCGGAGGTGTTCCGAGC 607
DB 572 ACTTTAGAAAATGGAACACGAGGCGCCCTGCTGCATATCTGAGAGGTTTCCAGAGC 631
QY 608 CCACTGTGGAGTGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGGCCCTG 667
DB 632 CGATCGTGAATGGTGTCTTGGATTCACAGGGGGAAGCTGTAAAGAAAGAGTCCAG 691
QY 668 CTGTTGTCAGAAAGGAGGAAAGTACTTTCATGAGTTGTTCCGAAACAGATCAGATGCT 727
DB 692 CTGTTGTTAAAGAGGAGGAAAGTCTTCATCAATTTATTTGGACGACATAGGTGCT 751
QY 728 GTGCTAAGATGACTGGGCGCGAATGCACCAAGCTGTTCAACCATAGATCTAAACCAG 787
DB 752 GTGCCAGAAATGAATGGGAGGGAATGCACAGGCTGTTTCAACATAGATCTAATCAAA 811
QY 788 CTCTCAGACACACTGCCCGAGTATTCTGAAAGTGGGGAAACCTTGTGGATCAGT 847
DB 812 CTCTCAGACCACTTGCCACAATTTCTTAAAGTAGGGGAACCTTTATGGATAAGGT 871
QY 848 GTAAAGCCATCCATGTGAACCATGGATTCGGGCTCACTTGGAGCTGGAGACAAAGCCC 907
DB 872 GCAAGCTGTTCATGTGAACCATGGATTCGGGCTCACTTGGGAATTAAGAAACAAGCAC 931
QY 908 TGAAGAGGCGCAGTACTTTGAGATGAGTACCTACTCCACAACAGACCATGATTCGGA 967
DB 932 TCAGAGGGGCACTACTTTGAGATGAGTACCTATTCAACAACAGAACTATGATACGGA 991
QY 968 TTCTCTTGGCTTTGTGCTCTCGTGGGAAGGAACGACACCGGATATTACACCTGCTCTT 1027
DB 992 TTCTGTTGCTTTGTATCATCATGTCGCAAGAAACGACACCGGATATCATCTTGTCTCT 1051
QY 1028 CTTCAAGACACCCAGCAGCTGAGCTTGGTGGCCATCTCTAGAAAAAGGTTTATAACG 1087
DB 1052 CTTCAAGACATCCAGTCAATCAGCTTTGGTACCATCGTAGGAAGGATTTATAATG 1111
QY 1088 CTACCACTCGAAGAGAGATGAATTTGACCCGCTAGCAAAAGTCTTCTGCTTCAGTCA 1147
DB 1112 CTACCAATTAAGTGAAGATTAATAATGACCAATATGAAGAGTGTGTTTTTCTGTCA 1171
QY 1148 GGTTTAAGCGTACCCACGAATCCGATGACGTGGATCTCTCACAAGCTCATTTCTCTT 1207
DB 1172 GGTTTAAGCGTACCCACGAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCTCT 1231
QY 1208 GTGAACAGAGAGCCTCGAGGATGGGTACAGCATATCTAAATTTTGGGATCATTAAGACA 1267
DB 1232 GTGAGCAAAAGGCTTTGATAACGATACAGCATATCCAAGTTTGGCAATCATTAAGCAC 1291
QY 1268 AGCCAGSAGATACATATCTGACAGAAATGATGACGCCAGTTCACCAAAATGTTCA 1327
DB 1292 AGCCAGSAGAAATATATTTCCATGACAGAAATGATGATGCCAATTTACCAAAATGTTCA 1351
QY 1328 CGCTGAATATAAGAAACCACTCAAGTGTAGCAAAATGGCTCAGCCAGCCAGCGCTCT 1387

DB 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTCGAGAAGCATCGCAAGTCTCAGGCTCT 1411
QY 1388 GTTCTCTGTAGTGTACCCGCTACCCGCTACCCGCTGGAAAGTGTTCGGACAATCTC 1447
DB 1412 GTTCTCGGATGGATACCCATACCATCTTGAGCTTGAAGAAGTGTTCAGACAAGTCTC 1471
QY 1448 CCAATTCACAGGAGGAATCCAGAGGAGTTTGAATATAAAGGCTTAACAGAAAAGTGT 1507
DB 1472 CCAACTGCACAGAGAGATACAGAGAGTCTGGAATAGAAAGCTACAGAAAAGTGT 1531
QY 1508 TTGCCAGTGGGTCTCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGCTTCTG 1567
DB 1532 TTGCACAGTGGTCTCGAGCAGTACTCTAAACATGAGTGAAGCCATATAAAGGTTCTCTG 1591
QY 1568 TCAATGCTGTGCTACAAATCTATGGGCAGTCTTGGCAACGCTTTCGAAAACCATCTTTTAAACTCAC 1627
DB 1592 TCAAGTGTGTGCATACAAATCCCTTGGCACATCTTGTGAGACGATCTCTTTTAAACTCTC 1651
QY 1628 CAGCCCTCTCCCTTTCATCCAAAGACAACATCTCTCTTATCGCAACCTTTGGGCTCTGTC 1687
DB 1652 CAGCCCTCTCCCTTTCATCCAAAGACAACATCTCTCTTATGCAACAATTTGGTGTGTC 1711
QY 1688 TCCCTCTCATTTGTTCTCTCATTTGTTGATCGCCACAAATACAAAAGCAATTTAGT 1747
DB 1712 TCCTCTCATTTGCTTTTAAACCTGCTAAATTTGTCAAGTACAAAAGCAATTTAGT 1771
QY 1748 ACGAGTCACTGCAGATGATCCAGTGTGCTGGCCCTGGATACGAGTACTTCTAGC 1807
DB 1772 ATGAAAGCCAGTACAGATGTTACAGTGTGACCGCTCTCAGATATGAGTACTTCTAGC 1831
QY 1808 TTGACTTCAGGAGTATGAATATGATGATCAAAATGGAGTTTCCAGAGAAATTTAGAT 1867
DB 1832 TTGATTTTCAGAGAAATATGAATATGATCAAAATGGAGTTTCCAGAGAAATTTAGAT 1891
QY 1868 TTGGGAAGTCTGGGCTCTGGGCTGCGGAGGTGATGAACGCGCCAGGCTATGCGCA 1927
DB 1892 TTGGGAAGTCTAGGATCAGTGTCTTTGGAAAAGTGTGAACCAACAGCTTATGGA 1951
QY 1928 TTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGTGTAAAAGAGAAAGTGA 1987
DB 1952 TTAGCAAAACAGGAGTCTCAATTCAGTGTGCGCTCAAAATGCTGAAAGAAAGCAGACA 2011
QY 1988 GCTGTGAAAAGAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
DB 2012 GCTCTGAAGAGAGGAGTCTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
QY 2048 ACAACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGGAGTGTACTTGAATTTG 2107
DB 2072 AGAATATGTGAACCTGCTGGGGGCTGCACACTGTCAGGACCAATTTACTTGAATTTG 2131
QY 2108 AATATGTTGCTATGGTGAACCTCTCACTACCTAAGAAAGTAAAAGAGAGAGTTTCA 2167
DB 2132 AATATGTTGCTATGGTGAACCTCTCACTACCTAAGAAAGTAAAAGAGAGAGTTTCA 2191
QY 2168 GGACATGACAGAGATTTTGAAGACATATTTTCAAGTCTTACCTACTTCCAGGAC 2227
DB 2192 GGACTTGGACAGAGATTTTCAAGGAACACATTTTCAAGTCTTACCCACTTCCCAATCAC 2251
QY 2228 ATTCAAAATTCAGCATGCTGTTTCCAGAGAGTTCAGTTACACCGCCGCTTGGATCAGC 2287
DB 2252 ATCCAAATTCAGCATGCTGTTTCCAGAGAGTTCAGATACACCGGACTCGGATCAAA 2311
QY 2288 TCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAAGAACAG 2347
DB 2312 TCTCAGGTTCAATGGGAATTCATTTCTGAGATGAGATTTGAATATGAAGAACAG 2371
QY 2348 AGAGCTGGCAGAGAGAGGAGGAGGATTTCAACGCTGCTGAGCTTTGAGACCTCTT 2407
DB 2372 AAAGGCTG-----GAAGAGAGGAGGAGTGTGAATGTGCTTACATTTGAAGATCTTCTT 2425
QY 2408 GCTTGTGCTACCAAGTGGCCAAAGCATGGAAATCTTGGAGTTCAAGTCTGTTGTCACCA 2467

Qy 728 GTGCTAGAAATGCACTGGCGCGGAATGCAACCAAGCTGTTCCACCATAGATCTAAACACAGG 787
Db 752 GTGCCAGAAATGAATCTGGCGAGGAATGCACACAGCTGTTCCAAATAGATCTAAATCAAA 811
Qy 788 CTCCTCAGAGCACACTGCCCCAGTTATTCTTGAAGTGGGGGAACCCCTTGGGATCAGGT 847
Db 812 CTCCTCAGAGCACATTCACCAATATTCTTAAAGTAGGGGAACCCCTTATGGATAAGGT 871
Qy 848 GTAAGGCCATCCATGTGAACCATGGATTGGGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907
Db 872 GCAAAGCTGTTCATGTGAACCATGGATTGGGGCTCACCTGGGGAATTAGAAACAAAGCAC 931
Qy 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAAGAGGACCATGATTCCGA 967
Db 932 TCGAGGAGGCGAATCTTTGAGATGAGTACCTATTCAACAAAGAGGATGATGATACCGA 991
Qy 968 TTCTCTTGGCCCTTGTCTTCCGTGGGAAGGAACGACACCGGATATTACACCTGCTCTT 1027
Db 992 TTCTGTTGGCTTTGTATCATCAGTGGCAAGAACGACACCGGATATTACACTTGTCTCT 1051
Qy 1028 CCTCAAGACCCCGACGAGTCAGCGTTGGTGACCATCTAGAAAGGGTTTATAACG 1087
Db 1052 CTTCAAGCATCCCGATCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAAATG 1111
Qy 1088 CTACCAAGTCGCAAGAGAGTATGAATTTGACCGGTAGCAAGGTTCTCTCAGTCA 1147
Db 1112 CTACCAATTTCAAGTGAAGATTATGAATTTGACCAATATGAAGGTTTGTCTCTGTC 1171
Qy 1148 GGTTTAAAGCTACCCAGCAATCGATCGAGTGGATCTCTCAAGCCCTCATTTCCCT 1207
Db 1172 GGTTTAAAGCTACCCACAATCAGATGTACGTGGACCTTCTCGAATCATTTCCCT 1231
Qy 1208 GTGAACAGAGAGGCTCGAGGATGGGTACACATATCTAAATTTGGCATATAGAACA 1267
Db 1232 GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAGTTTGTGCAATCATAGCAC 1291
Qy 1268 AGCCAGAGAGTACATATCTATCAGAAAATGATGACCGCCAGTTCAACCAAAATGTTCA 1327
Db 1292 AGCCAGAGAAATATATTTCCATCAGAAAATGATGATGCCCAATTTACCAAAATGTTCA 1351
Qy 1328 CGCTGAATATAAGAAAGAACTCAAGTCTAGCAAAATGCCTCAGCCAGCCAGGGCTCT 1387
Db 1352 CGCTGAATATAAGAAAGAACTCAAGTCTCAGAGAGCATCGCAAGTCAGGCTCT 1411
Qy 1388 GTTCCCTCTGATGCTACCCGTACCCCTCTTGACCTTGGAGTGGAAAGTGTTCGGACAAATCTC 1447
Db 1412 GTTCTCGGATGGATACCCATTACCATCTTGGACCTTGGAAAGTGTTCAGACAAAGTCTC 1471
Qy 1448 CCAATTCGACGAGGAATCCAGAAAGAGTTTGGAAATAAAAGGCTAACAGAAAAGTGT 1507
Db 1472 CCAACTGCAGAGAGATACAGAAAGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT 1531
Qy 1508 TTGGCCAGTGGGTCTCGAGCAGTACTCTAAATATGATGAGGCGCCGGGAAAGGCTCTGG 1567
Db 1532 TTGGACAGTGGGTCTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGTTCCCTGG 1591
Qy 1568 TCAAAATGCTGTGCGTACAAATTTCTATGGGACGCTTGGGAAACCATCTTTTAAACTCAC 1627
Db 1592 TCAAGTGTGTGCATCAAAATTTCCCTTGGCATATCTTGTGAGACGATCCCTTTTAAACTCTC 1651
Qy 1628 CAGGCCCCCTTCCCTTTCATCCAGACACATCTCTCTATGTCGACCATTTGGGCTCTGTC 1687
Db 1652 CAGGCCCCCTTCCCTTTCATCCAGACACATCTCTCTATGTCGACCAATTTGGTGTGTC 1711
Qy 1688 TCCCTCTCATTTGTTCTCTCATGTTGATGCTGCCACAAATACAAAAGCAATTTAGGT 1747
Db 1712 TCCTCTCATTTGCTGTTTACCCCTGCTAAATTTGTACAAAGTACAAAAGCAATTTAGGT 1771
Qy 1748 ACAGAGTACAGTGCAGATGATCCAGGTGACTGGCCCTCGATTAACAGGATCTTCTACG 1807
Db 1772 ATGAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCAGATAATGATGACTTCTACG 1831

Qy 1808 TTGACTTTCAGGAGCTATGAATATGACCTTAAGTGGGAGTTCGCGAGAGAACTTAGAGT 1867
Db 1832 TTGATTTTCAGAGCAATATGAATATGATCTCAATATGGAGTTTCCAAGAGAAAATTTAGAGT 1891
Qy 1868 TTGGAAAGTCTCTGGGCTCTGGCGTTCGCGAGGTGATGAACGCCACGGCTATAGGCA 1927
Db 1892 TTGGAAAGTCTTAGGATCAGGTGCTTTTGGAAAAGTGAACGCAACAGCTTATGGAA 1951
Qy 1928 TTAGTAAACGGAGCTCTCAATTCAGGTGCGGTGAAGATGCTTAAAGAGAAAGCTCACA 1987
Db 1952 TTAGCAAAACAGAGTCTCAATCCAGGTTGCGCTCAAAATGCTGAAGAAAACAGACACA 2011
Qy 1988 GCTGTGAAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
Db 2012 GCTCTGAAGAGAGGCACTCATGTCAAGACTCAAGATGATGACCCAGTGGGAAGCCACG 2071
Qy 2048 ACAACATCGTGAATCTGTGGGGGCAATGCACACTGTCAAGGCGCAGTGTACTGATTTTGT 2107
Db 2072 AGAATATTGTGAACCTGCTGGGGGCGTGCACACTGTCAAGACCAATTTACTTGTATTTTG 2131
Qy 2108 AATATTGTTGCTATGTTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAAAGTTTTCACA 2167
Db 2132 AATACTGTTGCTATGTTGATCTTCACTATCTAAGAAAGTAAAGAGAAAATTTTCACA 2191
Qy 2168 GGACATGGACAGAGATTTTAAAGGAACATAATTTCACTTCTTACCCTACTTTTCCAGCAC 2227
Db 2192 GGACTTGGACAGAGATTTCAAGGAACACAAATTTCACTTTTACCCCACTTTTCCAAATCAC 2251
Qy 2228 ATTCAAAATTCAGCATGCTGTTTTCAGAGAAAGTTCAGTTTACCCGCCCTTGGATCAGC 2287
Db 2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAAAGTTCAGATACACCCGGACTCGGATCAAA 2311
Qy 2288 TCTCAGGTTTCAATGGGAATTCAAATTCATCTCAAGATGAGATGAAATGAAAACAGA 2347
Db 2312 TCTCAGGCTTCAATGGGAATTCATTTCACTCTGAAGATGAAATGAAATGAAAACAAA 2371
Qy 2348 AGAGGCTGCGCAAGAAAGAGAGGAAAGATTGCAACGTGCTGAGCTTTGAAGAGCTCTCTT 2407
Db 2372 AAAGGCTG-----GAAGAGAGAGGAGGACTTGAATGCTGTACATTTGAAGATCTTCTT 2425
Qy 2408 GCTTTCGTTACCAAGTGGCCAAAGCATGGAATTCCTCGGAGTTCAGTCTGTTGTCACA 2467
Db 2426 GCTTTCGATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTGTTTCA 2485
Qy 2468 GAGACCTGCGACCCAGGAATGTGTTGTCACCCACCGGAAGGTGGTGAAGATCTGTGACT 2527
Db 2486 GAGACCTGGCCCGCAGGAACGCTGTTGTCAACCACCGGAAAGTGGTGAAGATGATGACT 2545
Qy 2528 TTGGAGCTGGCCGAGACATCTTGAGCGACTTCCAGCTACGTCGAGGGGCAAGCCAGCGC 2587
Db 2546 TTGGATTTGCTCGAGATATCATGAGTATTCCAACTATGTTGTGAGGGCAATGCCCCGTC 2605
Qy 2588 TGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACAAATCAAGAGTG 2647
Db 2606 TGCCCTGTAATATGGATGGCCCCGAAAGCTGTTTGAAGGCATCTACACCATTAAGAGTG 2665
Qy 2648 AGCTGTGGTCTACGSCATCTCTCTGGAGATATTTTCACTGGGTGTCGAACCTTACC 2707
Db 2666 ATGCTGTGTCATATGGAATTTACTGTGGAAATCTTCTCACTGGGTGTAATCTTACC 2725
Qy 2708 CTGGCATCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC 2767
Db 2726 CTGGCATCTCCGTTGATGCTACTTCTACAACTGATTCAAATGATTTAAATGATC 2785
Qy 2768 AGCCATTTCTATGCCAGAGGGATATCTTTGTAATGCAATCTGCTGGGCTTTTGAAT 2827
Db 2786 AGCCATTTTATGCTACAGAAAGAAATATACATTAATGCAATCTGCTGGGCTTTTGAAT 2845
Qy 2828 CAAGAGGCGGCATCTTCCCAACCTGACTTCATTTTATAGATGTCACGCTGCCAGAG 2887
Db 2846 CAAGGAAAGCGGCATCTTCCCTTAATTTGACTTCGTTTTTATAGGATGTCAGCTGGCAGATG 2905
Qy 2888 CAGAAGAAG-----CATGTATCAGAACATCCCATCTTACC 2924

```

Db      2906 CAGAGAGCGATGATCAGAAATCGGATGGCGGTGTTTCGGAATGCTCTCACACTACC 2965
          |||||
Qy      2925 AAAACAGCGCGCCCTCAGCAGAGAGGGGGCTCAGAGC-CAGATCGCCACAGCGCCAGG 2983
          |||||
Db      2966 AAAACAGCGCGACCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCGCAGGGCTCAGG 3025
          |||||
Qy      2984 T-GAAGATTACAGAGAAAGAGTTAGCAGAGAGCGCTTGGACCCGCC- ---ACCCTAG 3038
          |||||
Db      3026 TCGAAGATTCTGAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
          |||||
Qy      3039 CAGGCTGTAGACCGCAGAGCCCAAGATTAGCCTCGCCTCT- -GAGGAAGCGCCCTACAGCG 3096
          |||||
Db      3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTCATCACTAAAGAAATCTATTATCAAC 3145
          |||||
Qy      3097 CGTTCCTCGCGGACTTTCTCTAGATGCTGCTCGCCATTACTC- ------CAAGTGA 3149
          |||||
Db      3146 TGCTGCTTCACACAGACTTTCTCTAGAGCCGCTCGCGTTTACTCTGTTTCAAAGGGA 3205
          |||||
Qy      3150 CTTCTATAAAATCAAACTCTCTCGCACAGCGGAGAGCCAATAATGAGACTTGTGG 3209
          |||||
Db      3206 CTTTGTAAATCAAAATCATCTGTCACAAAGCAGAGAGCTGATAATGAATTTATTG 3265
          |||||
Qy      3210 TGAGCCCGCTACCCCTGGGGCTTTCCACGAGCTTGAGGGAAAGCCATGTATCTGA 3269
          |||||
Db      3266 GAGCATTGATCTGCATCCAAAGCCCTTCTCAGCGCGCTTGAGTGAATGTGTACCTGAAG 3325
          |||||
Qy      3270 TATAGTATATCTTCTGTAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGAAAGC 3329
          |||||
Db      3326 TACAGTATATCTTCTGTAATACATAAAACAAA- ------AGCATTTTGTAAAGGAAAGC 3378
          |||||
Qy      3330 TAAATATGATTTTAAATATCTATGTTTAAATACTATGTAACATTTTTCATCTATTAG 3389
          |||||
Db      3379 TAATATGATTTTTT- -AAGTCTATGTTTAAATATAATGTAATTTTTCAGCTATTAG 3436
          |||||
Qy      3390 TCATATATTTTATGATGGAATAAACTTTCTACTGTATAAAATAAAATAAAATAAAATAAA 3449
          |||||
Db      3437 TGATATATTTATGGTGGGATATAAAATTTCTACTACAGAAATAAAATAAAATAAAATAAA 3496
          |||||
Qy      3450 AAAA 3453
          |||||
Db      3497 AAAA 3500

RESULT 10
LOCUS      125170
DEFINITION Sequence 3 from patent US 5548065.
ACCESSION 125170
VERSION    125170.1 GI:1605040
KEYWORDS
SOURCE
ORGANISM   Unknown.
            Unclassified.
REFERENCE  1 (bases 1 to 3501)
AUTHORS   Lemischka,I.R.
TITLE      Tyrosine kinase receptor human flk-2-specific antibodies
JOURNAL    Patent: US 5548065-A 3 20-AUG-1996;
FEATURES   Location/Qualifiers
            source
              1..3501
              /organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 64.5%; Score 2227.2; DB 6; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy      8 GCCTACCGCGCGCTCGGAGGCCATCGGGCGGTGGCGAGCGCAGCGACCGCGCGCTGC 67
          |||||
Db      35 GGGGACCCCGGCTCGGAGGCCATCGCGCGCTTGGC- ---GCGCAGCGGGCACCGTGC 91
          |||||
Qy      58 TGCTGCTGTTGTTTGTGTCAGTAATGATTCCTTGAGACCGTTACAAACCAAGACCTGCCTG 127
          |||||
```

```

Db      92 CGCTGCTCCTGTTTCTGCAATGATTTGGGACTATTACAAATCAGATCTGCCTG 151
          |||||
Qy      128 TGATCAAGTGTGTTTAAATCAGTCAATGAGAAATGGCTCATCAGGGGAAAGCCATCAT 187
          |||||
Db      152 TGATCAAGTGTGTTTAAATCAATCAATGAGAAATGATTCATCAGTGGGGAAGTCATCAT 211
          |||||
Qy      188 CGTACCGAATGTGCGAGGATCCCCAGAAAGACCTCCAGTGTACCCCGAGCGGCAGAGTG 247
          |||||
Db      212 CATATCCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTGGTGTGAGACCCAGAGCT 271
          |||||
Qy      248 AAGGGACGCTATATGAGCGGCGCACCGTGGAGTGGCGGAGTCTGGGTCCATCACCCCTGC 307
          |||||
Db      272 CAGGGACAGCTGTACGAAGCTGCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331
          |||||
Qy      308 AAGTGCAGCTCGCCACCCAGCGGAGACCTTTTCCTGCTCTGGGTCTTTAAGCAGAGCTCCC 367
          |||||
Db      332 AAGTGTGCTGATGATGCCAGGGAACATTTCTCTGCTCTGGGTCTTTAAGCAGAGCTCCC 391
          |||||
Qy      368 TGGGCTGCCAGCGGCACTTTTGATTTACAAACAGAGGAATCGTTTCCATGCCATCTTGA 427
          |||||
Db      392 TGAATTCGCCAGCCACATTTTGATTTACAAACAGAGAGTGTGTTTCCATGGTCATTTTGA 451
          |||||
Qy      428 AGTGTACAGAGACCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGGAACGCCCAACT 487
          |||||
Db      452 AAATGACACAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGTACCAATT 511
          |||||
Qy      488 ACACAGTACTGTTTCACAGTGAATGTAAAGATACACAGCTGTATGTGCTTAAGAGACCTT 547
          |||||
Db      512 ACACAAATATTTTACAGTGAGTAAAGAAATACCCCTGCTTTACACATTAAGAGACCTT 571
          |||||
Qy      548 ACTTTAGGAGATGGAAAAACCCAGGATGCACCTGCTGTCATCTCCGAGGGTGTTCGGAGC 607
          |||||
Db      572 ACTTTAGAAAAATGGAACACCCAGGACGCCCTGCTGTCATATCTGAGAGCGTTTCCAGAGC 631
          |||||
Qy      608 CCACCTGTGAGTGGGTGCTCTGCAGCTCCACAGGGAAGCTCTAAAGAAAGAGGCCCTG 667
          |||||
Db      632 CGATCTGTGAATGGGTGCTTTGGGATTCACAGGGGGAAGCTGTAAAGAGAAAGTCCAG 691
          |||||
Qy      668 CTGTTCTCAGAAAGAGGAAAAAGGTACTTCATCAGTTGTTTCGGAACAGACATCAGATGCT 727
          |||||
Db      692 CTGTTCTTAAAAAGGAGGAAAAAGTCTTCATCAATATTTTGGGACGCGATAGGTGCT 751
          |||||
Qy      728 GTGCTAGAAATGCACCTGGGCGGGAATGCACCAAGCTGTTCCACATAGATCTAAACAGG 787
          |||||
Db      752 GTGCCAGAAATGAACCTGGCGAGGAATGCACCAAGGCTGTTCACAATAGATCTAAATCAA 811
          |||||
Qy      788 CTCCTCAGAGCACACTGCCCGCAGTTATTCTGAAAGTGGGGGAACCCCTTGTGATCAGGT 847
          |||||
Db      812 CTCCTCAGACACACTGCCACAAATTTCTTAAAGTAGGGGAACCCCTTATGGATAAGT 871
          |||||
Qy      848 GTAAGCCCATCCATGTGAACCATGGATTCGGGTCCAGCTCGGAGCTGGAAGACAAAGCCC 907
          |||||
Db      872 GCAAGACTGTTTCATGTGAACCATGGATTCGGGTCCAGCTGGGAATTAGAAAAACAAGCAC 931
          |||||
Qy      908 TGGAGGAGGCGCAGCTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 967
          |||||
Db      932 TCGAGGAGGCGCACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 991
          |||||
Qy      968 TTTCTTTGGCTTTGTTGCTTTCGTTGGAGGAACGACACCCGGATATTACACCTGCTCT 1027
          |||||
Db      992 TTTCTTTGCTTTTGTATCATCATGTCGCAAGAACGACACCCGGATATTACACTTGTCTCT 1051
          |||||
Qy      1028 CCTCAAGCACCCCGCAGCTCAGCTGTTGGTGACCATCTCTAGAAAAAGGTTTATAACG 1087
          |||||
Db      1052 CTTCAAGCATCCAGCTCAATCAGCTTGGTTTACCATCGTAGGAAGAGGATTTATAAGT 1111
          |||||
Qy      1088 CTACAGCTCGCAAGAGAGATGAAATTTGACCCGCTACGAAAGTTCGCTTCTCAGTCA 1147
          |||||
Db      1112 CTACCAATTCAAGTGAAGATTTATGAATTTGACCAATATGAAGAGTTTGTGTTTCTCTCA 1171
          |||||
Qy      1148 GGTTTAAGCGTACCACGAATCCGATGCACTGGATCTTCTCTCAAGCTCATTTCTCTT 1207
          |||||
```


Db 1172 GTTTTAAAGCCTACCCACAATAATCAGATGATCGTGGACCTCTCTCGAAAAATCATTTTCCTT 1231
Qy 1208 GTGACAGAGAGGCTGGAGATGGGTACACATATCTAAATTTTGGCATCATAGACA 1267
Db 1232 GTGAGCAAAAGGTTGATTAACGGATACAGCATATCCAAGTTTGGCAATCATAGCAC 1291
Qy 1268 AGCCAGGAGATACATATTTCTATCAGAAAAATGATGACGCCAGTTTACCACCAAAATGTTCA 1327
Db 1292 AGCCAGGAGATATATTTCCATCGAANAATGATGATGCCAATTTACCAAAATGTTCA 1351
Qy 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAGGCGTCT 1387
Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGAGAAGCATCGGCAAGTCAGGCGTCT 1411
Qy 1388 GTTCTCTGATGGCTACCCGTACCTCTTGGACCTGGAAGAGTGTTCGACAAATCTC 1447
Db 1412 GTTCTCGGATGATACCCATTTAGCATCTTGGACCTGGAAGAAGTGTTCGACAAATCTC 1471
Qy 1448 CCAATTGCAGGAGAAATCCAGAAAGAGTTTGGAAATAAAGGCTAACAGAAAGTGT 1507
Db 1472 CCACTGCACAGAGAGATCACAGAGAGTCTGGAATAGAAAGCTAACAGAAAGTGT 1531
Qy 1508 TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAAGGCTTCTGG 1567
Db 1532 TTGGACAGTGGGTGTGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGTTTCTGG 1591
Qy 1568 TCAATGCTGTGCTACAAATTTATGGCCAGCTTTGGCAACCATCTTTTAACTCAC 1627
Db 1592 TCAAGTGTGCTGATACAAATTTCCCTTGGCAGCATCTTGTGAGACGATCTCTTAAACTCTC 1651
Qy 1628 CAGGCCCTTCCCTTTTATCCAAGACAATCTCTCTCTATGCGACCATTTGGGCTCTGTC 1687
Db 1652 CAGGCCCTTCCCTTTTATCCAAGACAATCTCTCTCTATGCGACAATTTGGTGTGTC 1711
Qy 1688 TCCCTCTCATTTGTTCTCATTTGTGTGATCTGCCCAAAATACAAAAGCAATTTAGGT 1747
Db 1712 TCCTCTTCAATGTGTTTTAAACCTGCTAAATTTGTCAAGTACAAAAGCAATTTAGGT 1771
Qy 1748 ACGAGATCACTGCAGATGATCCAGTGTACTGGCCCTGGATTAACAGTACTTCTAG 1807
Db 1772 ATGAAGCCAGCTACAGATGTACAGTGTGACGGCTCTCTCAGATAATGAGTACTTCTACG 1831
Qy 1808 TTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAACTTAGAGT 1867
Db 1832 TTGATTCAGAGATATGAATATGATCTCAATGGAGTTTCCAGAGAAATTTAGAGT 1891
Qy 1868 TTGGGAAGTCTCGGGTCTCGCGCTTTGGGAGGTTGATGAACGCCACGCGCTATGGCA 1927
Db 1892 TTGGGAAGTACTAGGATCAGTGTGCTTTTGGAAAAGTGTATGAACGCAACAGCTTATGGAA 1951
Qy 1928 TTAGTAAACGGGAGTCTCAATTCAGTGTGGGTGAAGATGCTAAAGAGAAAGTGCAC 1987
Db 1952 TTAGCAAAACAGGAGTCTCAATTCAGTGTGGGTCAAAATGCTGAAGAAAAGAGAGACA 2011
Qy 1988 GCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACCACTAG 2047
Db 2012 GCTCTGAAGAGAGGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
Qy 2048 ACAACATCGTAATCTGTGGGGGATGCACACTGTGAGGGCCAGTGTACTTTGATTTTG 2107
Db 2072 AGAATATTGTGAACCTGTGGGGGCTGCACACTGTGAGGACCAATTTACTTTGATTTTG 2131
Qy 2108 AATATTGTTGATGGTGACCTCTCACTACTAGTAAAGTAAAAGAGAAAGTGTTCACA 2167
Db 2132 AATACTGTTGATGGTATCTCTCACTATCTAAGAGTAAAAGAGAAAAATTTTCACA 2191
Qy 2168 GGACATGCACAGAGATTTTAAAGACATAATTTAGTCTTCTTACCTACTTTCCAGGCAC 2227
Db 2192 GGACTTGGACAGAGATTTTCAAGGAACACAATTTTCAAGTCTTTTACCCCACTTTTCCATCAC 2251
Qy 2228 ATTCAAATTCAGAGTGTGTTTCCAGGAGAGTTCAGTTACACCCGCGCTTTGGATCAGC 2287
Db 2252 ATCCAAATTCAGAGTGTGTTTCAAGAGAAAGTTCAGATACACCCGCGACTCGGATCAA 2311

Qy 2288 TCTCAGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATGAATAAACAGA 2347
Db 2312 TCTCAGGCTTCATGGGAATTCATTCCTACTCTGAAGATGAATGAATGAATAAACAAA 2371
Qy 2348 AGAGCTGGCAGAGAAGAGAGAGAGATTTGAACGTGTGAGCTTTGAAGACCTCTCTT 2407
Db 2372 AAAGCTG-----GAAGAAGAGGAGGACTTGAATGTCTTACATTTGAAGATCTCTT 2425
Qy 2408 GCTTTGCTGCTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGTGTCACA 2467
Db 2426 GCTTTGCTATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2485
Qy 2468 GAGACTGGCAGCAGAGAAATGTTGGTCAACCACGGGAAGTGTGAAGATCTGTGACT 2527
Db 2486 GAGACTGGCCGCGAGGAAGTGTCTGTCAACCACGGGAAGTGTGAAGATATGACT 2545
Qy 2528 TTGACTGGCCGCGAGACATCTGTGAGCAGTCCAGCTACGTGTCAGGGGCAACGACGGC 2587
Db 2546 TTGGATTTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTGTCAAGGGCAATGCCGTC 2605
Qy 2588 TGCCGCTGNAAGTGGATGGCCACCGAGAGCTTATTTGAAGGATCTACACATCAAGAGTG 2647
Db 2606 TGCCCTGTAANAATGGATGGCCCGGAAAGCTGTTTGAAGGCATCTACACATTAAGAGTG 2665
Qy 2648 AGTCTGGTCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
Db 2666 ATGCTGTGTCATATGAATATTACTGTGGAAATCTTCTCACTTTGGTGAATCTTACC 2725
Qy 2708 CTGGCATTCCTGTGCGACGCTAACTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC 2767
Db 2726 CTGGCATTCCTGGTGTGATGCTAACTCTACAACTGATTTCAAAATGGATTTAAATGGATC 2785
Qy 2768 ACCATTTCTATGCCAGAGAGATATCTTTGTATGCAATCTCTGCTGGGCTTTTGGACT 2827
Db 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCTCTGCTGGGCTTTTGGACT 2845
Qy 2828 CAAGGAAGCGGCATCTTCCCACTGACTTCAATTTTGGATGTGACGTGGCAGAGG 2887
Db 2846 CAAGGAAGCGGCATCTTCCCTTAATTTGACTTCGTTTTTGGATGTGACGTGGCAGATG 2905
Qy 2888 CAGAAGAAG-----CATGTATCAGAACATCCATCCATCTACC 2924
Db 2906 CAGAAGAAGCGATGTATCAGATGTGATGGCGGTGTTTCCGAATGTCTTCAACACTACC 2965
Qy 2925 AAAACAGGCGGCTTCAAGAGAGAGCGGCTCAGAGC--CCAGTCCCGACAGCGCCAGG 2983
Db 2966 AAAACAGGCGGCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025
Qy 2984 T-GAAGATTACAGAGAAAGAAAGTGTAGCGAGGAGGCTTGGACCCCGCC-----ACCCCTAG 3038
Db 3026 TCGAAGATTCTAGAGAGACAATTTAGTTTAAAGACTTCTATCCCTCCACTATCCCTAA 3085
Qy 3039 CAGGCTGTAGCGCAGAGAGCCAAAGATAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGTTTACCAAAACAAGATTAATTTATCATCACTAAAAGAAATCTATTATCAAC 3145
Qy 3097 CGTTGCTTGGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA 3149
Db 3146 TGTGCTTCCACAGACTTTTCTTAGAAGCGCTGCTGTTTACTTCTTTTCAAGAGGA 3205
Qy 3150 CTTCATAAATAACAACTCTCTCTCGACAGCGGAGAGCCCAATAATGAGACTTGTGG 3209
Db 3206 CTTTGTGAATTAATAATCATCTCTGTCACAGGAGGAGGAGCTGATTAATGACTTTATTG 3265
Qy 3210 TGAGCCCGCTACCCCTGGGGCTTTTCCAGAGCTTTGAGGGGAAAGCATGTATCTGAAA 3269
Db 3266 GAGCATTTGATCTGCATCCRAAGGCTTCTCAGCGCGGCTTGAGTGAATTTGTACTCTGAAG 3325
Qy 3270 TATAGTATATCTTTGTAATATACGTGAACCAAAACCAACCCGTTTTCGTTGAAGGAAGC 3329
Db 3326 TACAGTATATCTTTGTAATATACATAAAACAAA-----AGCATTTTTCGTAAGGAGAGC 3378

Qy	3330	TAAATATGATTTTAAAAAATCTATGTTTAAAAATACATATGTAACCTTTTTCATCTATTTAG	3386
Db	3379	TAATATGATTTT - AAGTCTATGTTTAAATAAATATGTAATTTTTCAGCTATTTAG	3436
Qy	3390	TGATATATTTATGGATGGAAATAAACTTCTACTGTAAAAAATAAAAAAAAAAAAAA	3449
Db	3437	TGATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAAAAAAAA	3496
Qy	3450	AAAA 3453	
Db	3497	AAAA 3500	
RESULT 11			
140601			
LOCUS	I40601	3501 bp	DNA 1 linear PAT 13-MAY-1997
DEFINITION	Sequence 3 from patent US 5621090.		
ACCESSION	I40601		
VERSION	I40601.1 GI:2082893		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 3501)		
TITLE	Lemischka, I.R.		
JOURNAL	Nucleic acids encoding soluble human FLK-2 extracellular domain		
FEATURES	Patent: US 5621090-A 3 15-APR-1997;		
source	Location/Qualifiers		
	1..3501		
BASE COUNT	1068 a	709 c	784 g
ORIGIN	940 t		
Query Match 64.5%; Score 2227.2; DB 6; Length 3501;			
Best Local Similarity 79.6%; Pred No. 0;			
Matches 277; Conservative 0; Mismatches 653; Indels 56; Gaps 10;			
Qy	8	GGCTACCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGCAGCGCGGCTGC	67
Db	35	GGGACCCCGGCTCCGGAGGCCATCGCGCTTGGC--GCGGACGCGGGCACCGCTGC	91
Qy	68	TGCTGCTTGTGTTTGTCTCAGTAATGATCTTGAGACCGTTACAAACACAGACCTGCCTG	127
Db	92	CGCTGCTCGTGTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG	151
Qy	128	TGATCAAGTGTGTTTAAATACAGTCATCAGAACATGGCTCATCAGCGGGAAGCCATCAT	187
Db	152	TGATCAAGTGTGTTTAAATCAATCATAAGAACAAATGATTCATCAGTGGGAAATCATCAT	211
Qy	188	CGTACCAATGGTCCGAGGATCCCCGAAGACCTCCAGTGATCCCCGAGCGCCAGAGTG	247
Db	212	CATATCCCATGGTATCAGATCCCGCGAAGACCTCGGTGTGCTTCGAGACCCAGAGCT	271
Qy	248	AAGGACGGTATATGAACGCCACCGTGGAGTGGCGGAGTCTGGGTCCATCACCTGC	307
Db	272	CAGGACAGTGTACGAAGCTGCCCTGTGGAAGTGGATGTATGCTTCCATCACACTGC	331
Qy	308	AAGTGCAGCTCGCCACCCAGGGGACCTTCCCTGCCCTGGTCTTTAAGCACAGCTCCC	367
Db	332	AAGTGTGCTGGATGCCCCAGGGAACATTCCTCTCTGGGTCTTTAAGCACAGCTCCC	391
Qy	368	TGGCTGCCAGCCGCACTTTGATTACAAAACAGAGGAATCGTTTCCATGGCCATCTTGA	427
Db	392	TGAATTCGCCACCAATTTGATTTACAAACAGAGGAGTGTGTTCCATGTCATTTGA	451
Qy	428	ACGTCAGACAGACCCAGCGAGGAATACCTTACCATATTCAGAGCGAAGCGGCCAACT	487
Db	452	AAATGACAGAAACCAAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAATT	511
Qy	488	ACACAGTACTGTTCCAGCTGNACTAGAGNATACACAGCTGTATGTCTTAAGGAGACCTT	547
Db	512	ACACAAATATTGTTACAGTGAATATAAGAAATACCCCTGCTTTACACAAATTAAGAGACCTT	571

Qy	2108	AAATATTGTTGCTATGTTGACCTCCTCAACTACCTAAGAACTAAAGAGAGAAATTTTCACA	2167
Db	2132	AAATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAAGTAAAGAGAAAAATTTTCACA	2191
Qy	2168	GGACATGGACAGAGATTTTAAAGGAAACATAATTTCACTTCTTAACTCCTACTTTTCCAGGCAC	2227
Db	2192	GGACTTGGACAGAGATTTTCAAGGAAACAAATTTTCACTTTTACCACCACTTTTCCATCACC	2251
Qy	2228	ATTCAAATTCACGATGCCTGGTTTCACGAGAAAGTTTCAAGTTTACACCGGCCCTTGGATCAGC	2287
Db	2252	ATCCAAATTCACGATGCCTGGTTTCAAGAGAAAGTTTCAAGATACACCCGGACTCGGATCAAA	2311
Qy	2288	TCTCAGGGTTCAATGGGAATTTCAATTCATTTCTGAAAGATGAGATTTGAATATGAAAAACGAGA	2347
Db	2312	TCTCAGGGCTTCATGGGAATTTCAATTTTCACTCTGAAAGATGAAATTTGAATATGAAAAACCAA	2371
Qy	2348	AGAGCTGGCAGAGAAGAGAGAGAGGAGGAGATTTGAACGTGTCTGACGTTTGAAGACCTCCTTT	2407
Db	2372	AAAGGCTG-----GAAGAAGAGAGACTTGAATGTGCTTACATTTTGAAGATCTTCTTT	2425
Qy	2408	GCTTTGCGTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTTCAAGTCTGCTGTGCCACA	2467
Db	2426	GCTTTGCATATCAAGTTGCCAAAGGAATGAAATTTCTGGAAATTTAAGTCGTGTGTTTTCACA	2485
Qy	2468	GAGACCTGGCAGCCAGGAATGTGTTGGTACCACCGGGAAGGTGGTGAAGATCTGTGACT	2527
Db	2486	GAGACCTGGCGCGCAGGAAGCTGTTGTCAACCCACCGGAAAGTGTGTGAAGATATGTGACT	2545
Qy	2528	TTGACTTGGCCCGAGACATCCTCAGGCGACTCAGCTACGTCGTCAGGGGCAACCGCACGGC	2587
Db	2546	TTGGATTGGCTCAGATATCATGAGTGATTCCAACTATGTGTTCAGGGGCAATGCCCGTC	2605
Qy	2588	TGCCGCTGAAGTGGATGGCACCGCAGAGCTTATTTGAAAGGATCTACACAATCAAGAGTG	2647
Db	2606	TGCCTGTAAATGGATGGCCCCGAAAGCTGTTTGAAAGCATCTACACCATTTAAGAGTG	2665
Qy	2648	ACGTCGTGGCTACGGCATCCTTCTCTGGAGATATTTTTCACGTGGGTGTGAACCCCTTACC	2707
Db	2666	ATGCTGTGGTCATATGGAATATTAAGTGTGGAAATCTTCACTGTGTGTGAATTCCTTACC	2725
Qy	2708	CTGGCATTCCTGTCGACGCTAACTTCTATAAAGTGAATTCAGAGTGGATTTTAAATGGAGC	2767
Db	2726	CTGGCATTCGGTTGATGCTAACTTCTACAACTGATTCAAATGGATTTAAATGGATC	2785
Qy	2768	AGCCATCTATGCCACAGAAAGGATATACATTTTGTATGCAATCCTGCTGGGCTTTTGAAT	2827
Db	2786	AGCCATTTTATGTCACAGAAATAATACATTTATAATGCAATCCTGCTGGGCTTTTGAAT	2845
Qy	2828	CAAGGAGCGGCATCCTTCCCAACCTGACTTCATTTTATAGATGTACAGCTGGCAGAGG	2887
Db	2846	CAAGGAACGGCCATCCTCCCTAAATTGACTTTCGTTTATAGATGTACAGCTGGCAGATG	2905
Qy	2888	CAGAAGAAG-----CATGTATCAGAACATCCATCCATCTACC	2924
Db	2906	CAGAAGAAGCATGTATCAGAAATGTGATGGCCGTGTTTCGGAAATGTCTCTCAACCTACC	2965
Qy	2925	AAACAGCGCGCCCTCAGCAGAGAGCGGGCTCAGAGC--CAAGTGCACACAGCGCCAGG	2983
Db	2966	AAACAGCGGCACCTTTTTCAGCAGAGAGATGATTTGGGGCTTACTCTCTCCGACGGCTCAGG	3025
Qy	2984	T--CAAGATTCACAGAGAAAGCTTACGAGGAGGCTTCGGACCC---CGCCACCCCTAG	3038
Db	3026	TCGAAGATTCGTAGAGGAACAAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA	3085
Qy	3039	CAGGCTGTACAGCGCAGAGCAAGATTAGCCTCGCCCTCT--GAGGAAGCGCCCTACAGCG	3096
Db	3086	CAGGCTGTAGATTACCAAAACAGATTAATTTTATCACTAAAAGAAATCTATTATCAAC	3145
Qy	3097	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCGCAATTAATCT-----CAAGTGA	3149
Db	3146	TGCTGCTTACCAGACTTTTCTCTAGAAAGCGCTCTGCTTTTACTCTGTTTTCATGTTTTCAGGGA	3205

Qy	3150	CTTCTATAAAATCAAACTCTCTCCGACACAGGCGGGAGAGCCCAATATGAGACTTGTTGG	3209
Db	3206	CTTTTGTAAATCAAAATCATCTCTGTACAAAGCGAGGAGCTGTATATGAACCTTTATTG	3265
Qy	3210	TGAGCCCGCCTACCTCGGGGGCCTTTCCACGAGCTTGAGGGGAAAGCCATGTAATCTGAAA	3269
Db	3266	GAGCATTTGATCTGCATCCAGGCCCTCTCAGGCCGGCTTGAGTGAATGTGTACTCTGAG	3325
Qy	3270	TATAGTATATCTTGTGTAATACGTGAACCAAAACCAACCCGTTTTTTCGTAAAGGAAAGC	3329
Db	3326	TACAGTATATCTTGTAAATACATAAACAACAA-----AGCATTTTTGCTAAGGAGAAGC	3378
Qy	3330	TAAATATGATTTTAAAAAATCTATGTTTTTAAAAATACTATGTAACCTTTTTTCATCTATTAG	3389
Db	3379	TAAATATGATTTTTT--AAGTCTATGTTTTTAAAAATATATGTAATTTTTTCAGCTATTAG	3436
Qy	3390	TGATATATTTATGATGGAAATAAACTTTCTACTGTAAA	3429
Db	3437	TGATATATTTATGGGTGGGAATAAAATTTCTACTACAGA	3476
RESULT 13			
LOCUS I44733			
DEFINITION Sequence 3 from patent US 5635388.			
ACCESSION I44733			
VERSION I44733.1 GI:2469446			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unknown.			
REFERENCE			
1 (bases 1 to 3475)			
AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.			
TITLE Antisense antibodies against the flk2/flt3 receptor and uses thereof			
JOURNAL Patent: US 5635388-A 3 03-JUN-1997;			
FEATURES			
source 1..3475			
BASE COUNT 1042 a 709 c 784 g 940 t			
ORIGIN			
Query Match 63.8%; Score 2203; DB 6; Length 3475;			
Best Local Similarity 79.6%; Pred. No. 0;			
Matches 2749; Conservative 0; Mismatches 650; Indels 56; Gaps 10;			
Qy	8	GGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCGAGCGACCGCGCGCTGC	67
Db	35	GGGAGACCCGGGCTCCGGAGGCCATCGCGCGTTGGC---GCGCGACGCGGGACCGTGC	91
Qy	68	TGCTGCTGTGTTTTCTGCTAGTATGATCTTGAGACCGTTTACAAACCAAGACCTGCCTG	127
Db	92	CGCTGCTGCTGTTTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG	151
Qy	128	TGATCAAGTGTGTTTTTAATCAGTCTATGAGAACAAATGGCTCATCGCGGAAAGCCATCAT	187
Db	152	TGATCAAGTGTGTTTTTAATCAATCATAAACAATGATTCATCAGTGGGAAAGTCATCAT	211
Qy	188	CGTACCGAATGGTCGGAGATCCCGCAGAACCTCCAGTGTACCCCGAGGGCCAGAGTG	247
Db	212	CATATCCCATGGTATCAGAAATCCCGGAAAGACCTTCGGGTGTGAGACCCCGAGAGCT	271
Qy	248	AAGGACGGGTATATGAAGCGGCCACCTGGAGGTGGCGGAGTCTGGGTGCATCAACCTGC	307
Db	272	CAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC	331
Qy	308	AAGTGTAGCTCGCCACCCCGAGGGAACCTTTCTCGCTCTGGGTCTTTAAGCACAGCTCCC	367
Db	332	AAGTGTGGTGCATGCGCCCGAGGAACATTTCTCTCTCTGGTCTTTAAGCACAGCTCCC	391
Qy	368	TGGCTCGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGCCATCTTGA	427
Db	392	TGAATTCGCCAGCCACATTTTGATTTACAAACAGAGGAGTGTGTTTCCATGGTCATTTGA	451

QY	368	TGGGCTGCCAGCCGACTTTGATTTACAAAACAGAGGAATCGTTCCATGGCCACTTGA	427	QY	1448	CCAATTGCAGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGAAAAGTGT	1507
Db	392	TGAATGGCCAGCCACTTTGATTTACAAAACAGAGGAGTTTCCATGGCTCAATTTGA	451	Db	1472	CCAACCTGCACAGAGAGATCACAGAGGAGTCTGGAATAGAGAGGCTACAGAAAAGTGT	1531
QY	428	AGCTGACAGAGACCAGGAGAGGAATACCTACTCCATATTCAGAGCGAACCGCAACT	487	QY	1508	TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGCGCGGGAAGGCTCTCTGG	1567
Db	452	AAATGACAGAAACCCCAAGCTGGAGAATACCTACTTTTATTCAGAGTGAAGCTACCAAT	511	Db	1532	TTGGCAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAAGGTTCTCTGG	1591
QY	488	ACACAGTACTGTCACAGTGAATGTAAGAGATACACAGCTGATGTGCTAAGGAGACCTT	547	QY	1568	TCAATATGCTGCTGCTACAATCTATGCGCAGCTCTTTCGAAACCATCTTTTAACTCAC	1627
Db	512	ACACAAATATGTTTACAGTGAATGAAGAAATACCCCTGCTTTACACATTAAGAAGACCTT	571	Db	1592	TCAGTGTGTCATACAATCCCTTGGCAGATCTTGTGAGAGATCCCTTTTAACTCTC	1651
QY	548	ACTTTAGGAAGATGAAAACACAGGATGCACTGCTCTGCATCTCCGAGGAGTGTTCGGAGC	607	QY	1628	CAGGCCCTTCCCTTTTCATCCAAAGACACATCTCTTATGACACATTTGGCTCTGTC	1687
Db	572	ACTTTAGAAAATGAAAACACAGGAGCCCTGCTGCATATCTGAGAGCTTCCAGAGC	631	Db	1652	CAGGCCCTTCCCTTTTCATCCAAAGACACATCTCTTATGCACAATTTGGTGTGTC	1711
QY	608	CCACTGTGGAGTGTCTGCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTG	667	QY	1688	TCCCTTCATTTGTTGCTCATTTGTTGATCTGCGCACAAATACAAAAGCAATTTAGGT	1747
Db	632	CGATCGTAATGGGTGCTTCCGATTCACAGGGGAAAGCTGTAAAGAAAGAAAGTCCAG	691	Db	1712	TCTCTCTTCAATGCTGCTTTTAAACCTGCTAATTTGTCAAGTACAAAAGCAATTTAGGT	1771
QY	668	CTGTTGTCAGAAAGGAGAAAGGTACTTTCATGAGTTGTTCCGAACAGACATCAGATGCT	727	QY	1748	ACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATTAACGAGTACTTCTACG	1807
Db	692	CTGTTGTTAAAGAGGAGAAAGGTCTTCATGAATTTTGGGACGGACATPAAGGTGCT	751	Db	1772	ATGAAGCCAGCTACAGATGGTACAGGTACCGGCTCTCAGATATAGTACTTCTACG	1831
QY	728	GTGCTAGAAATGCATGGGCCCGGAATGCACCAAGCTGTTCAACATAGATCTAAACACAG	787	QY	1808	TTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCTCCGAGAGAGAACTTAGAGT	1867
Db	752	GTGCCAGAAATGNACTGGGCGAGGAATGCACACAGCTGTTCAATATAGATCTTAATCAAA	811	Db	1832	TTGATTCAGAGAAATGAATATGATCTCAATGGGAGTTTCCAGAGAGAAATTTAGAGT	1891
QY	788	CTCCTCAGAGCACTGCCCCAGTTTATTCCTGAAAGTGGGGAACCCCTTGTGGATCAGGT	847	QY	1868	TTGGAAAGTCTCGGGTCTGGCTTTCGGGAGGTGATGAACGCCACGGCTATAGGCA	1927
Db	812	CTCCTCAGACCACATGTCACAAATTTTCTTAAAGTAGGGGAACCCCTATGATGAAGGT	871	Db	1892	TTGGAAAGTACTAGGATCAGGTGCTTTTGGAAAAGTGAACGCAACAGCTTATGGAA	1951
QY	848	GTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGCAAAAGCCC	907	QY	1928	TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAGAGCTCACA	1987
Db	872	GCAAAAGCTGTATGTAACCATGATTCGGGCTCACCTGGGAATTAGAAAAACAAGCAC	931	Db	1952	TTAGCAAAACAGAGTCTCAATCCAGTTTGGCTCAAAATGCTGGAAGAAAAGAGACACA	2011
QY	908	TGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA	967	QY	1988	GTCTGAAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACTGGGACACCATG	2047
Db	932	TCGAGAGGGACACTTCTGAGATGAGTACCTATTCACAAACAGAACTATGATACGGA	991	Db	2012	GTCTGAAAGAGGCACTCATGTCAGAACTCAGATGATGACCCAGCTGGGAGCCACG	2071
QY	968	TTCTCTTGGCCCTTGTCTTCCGTGGGAAGGAGACACCCGATATTACACCTGCTCTT	1027	QY	2048	ACAACATCGTGAATCTGTGGGGCATGCACACTGTCAGGGCCAGTGTACTGATTTTGT	2107
Db	992	TTCTGTTTGTCTTCTATCATCAGTGGCAAGAAACGACACCCGATACACTTGTCTCT	1051	Db	2072	AGAATATTTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATTTACTTGATTTTG	2131
QY	1028	CCTAAAGACCCAGCAGCTCAGGTTGGTGACCATCTAGAAAAGGTTTAAAGC	1087	QY	2108	ATAATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAGTAAAGAGAGAACTTTCACA	2167
Db	1052	CTTCAAGCATCCCACTCAATCAGCTTTGGTTACCATCGTAGGAAGGATTTATAAATG	1111	Db	2132	ATACTGTTGCTATGTTGATCTTCTCACTATCTAAGAGTAAAGAGAGAAATTTTCACA	2191
QY	1088	CTACCAGCTCGCAAGAGATGAAATTGACCCGTACGAAAGTTCTGCTTCTCAGTCA	1147	QY	2168	GGACATGGACAGAGATTTTAAAGGAACATAATTCAGTTCTTACCCTACTTTCCAGGCAC	2227
Db	1112	CTACCAATTCAGTGAAGATATGAATTTGACCAATATGAAGATTTGTTTTCTCTGCA	1171	Db	2192	GGACTTGGACAGAGATTTCAAGGAACACAAATTCAGTTTACCCCACTTTCCCAATCAC	2251
QY	1148	GGTTTAAAGCTTACCCAGGAATCCGATGCAGCTGGATCTTCTCAAGCCTCATTTCTCTT	1207	QY	2228	ATTCAAAATCCAGCATGCTGTTTACAGAGAGTTCAGTTACACCCGCCCTTGGATCAGC	2287
Db	1172	GGTTTAAAGCCTTACCCACAAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCTCT	1231	Db	2252	ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGGACTCGGATCAAA	2311
QY	1208	GTGAACAGAGAGCCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCATAGAACA	1267	QY	2288	TCTCAGGTTCAATGGGAATTCAAATTCATCTGAAAGTGAAGATGATGAAACCCAGA	2347
Db	1232	GTGAGCAAAAGGCTTGTATACGGATACAGCATATCCAAAGTTTGGCAATCATAGACACC	1291	Db	2312	TCTCAGGCTTCAATGGGAATTCATTTCACTCTGAAGTGAATGAAATGAAATGAAACCCAAA	2371
QY	1268	AGCCAGGAGATACATATCTATGCAGAAATGATGACCCAGTTCACCAAAATGTTCAC	1327	QY	2348	ACAGGCTGCGACAGAAAGAGAGGAGGATTTGAACGTGCTGACGTTTGAAGACCTCTCTT	2407
Db	1292	AGCCAGGAGATATATATTCATGAGAAAATGATGATGCCCAATTTACCAAAATGTTCAC	1351	Db	2372	AAAGGCTG-----GAAGAGAGGAGGACTTGAATGCTGTACATTTGAGATCTCTCTT	2425
QY	1328	CGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCAGCCAGGCGTCT	1387	QY	2408	GCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTGTTGTCCACA	2467
Db	1352	CGCTGAATATAGAAGAAACCTCAAGTGTCTCGCAGAGCATCGGCAAGTCAAGGCGTCT	1411	Db	2426	GCTTTGCATATCAAGTTGCCAAGGAATGAAATTTCTGGAAATTTAAGTCTGTTGTTTCA	2485
QY	1388	GTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGCAAAATCTC	1447	QY	2468	GAGACTGCGACGCCAGGAATGTTGGTCAACCCACCGGAGGAGTGTGAAGATCTGTGACT	2527
Db	1412	GTCTCTCGATGATACCATTTACCATCTTGGACCTGGAAGAGTGTTCAGACAAAGTCTC	1471	Db	2486	GAGACTGCGCCGAGGAACGCTGCTTGTCAACCCACCGGAAAGTGTGTGAAGATATGACT	2545
				QY	2528	TTGAGCTGCCCCGAGACATCTCTGAGGACTCTCAGCTACGTCAGCGGCAAGCGACGGC	2587

216 AGACCTCCAGTGTACCCGAGCGCCAGAGTGAAGGAGCGGTATATGAAGCGCCACCGT 275
QY
183 AGACCTCGGTCTGCGTGTGAGACCCAGAGCTCAGGACAGTGTAGGACGTCGCTGT 242
Db
276 GGAGGTGGCCGAGTCTGGGTCCATCACCGTGAAGTGCAGCTCGCCACCCAGGGGACCT 335
QY
243 GGAAGTGGATGTATCTGCTTCATCACACTGCAAGTGTGTCGATGCCCCAGGGAACAT 302
Db
336 TTCCTGCCCTCGGTCTTTAAGCAGAGCTCCCTGGCTGCCAGCGCACTTTGATTACA 395
QY
303 TTCCTGCTCTGGGTCTTTAAGCAGAGCTCCCTGAAATGCCAGCCACATTTTGATTACA 362
Db
396 AAACAGAGGAATCGTTTCCATGGCCACTTTGAACGTGACAGAGACCCAGCGAGGAATA 455
QY
363 AAACAGAGGAGTTGTTCCATGTCATTTTGAATGACAGAAACCCAGCTGGAGATA 422
Db
456 CCTACTCCATATTAGAGGGAACGGCCCACTTACAGACTACTGTTTCACAGTGAATGAAG 515
QY
423 CCTACTTTTATTACAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAATGAAG 482
Db
516 AGATACAGAGCTGTATGCTGAAGAGACCTTACTTTAGGAAGATGGAACCCAGGATCC 575
QY
483 AATACCTCGCTTTACAAATTAAGAGACCTTACTTTTGAATAATGGAACCCAGGATCC 542
Db
576 ACTGCTCTGATCTCCGAGGTGTTCCGGAGCCACTGTGAGTGGGTGCTCTGCAAGTCC 635
QY
543 CTGCTCTGATATCTGAGAGCTTCCAGAGCGATCGTGGATGGGTCTTTGCGATTC 602
Db
636 CCACAGGGAAGCTGTAAAGAAAGAGGCCCTCGCTGTTGTCAGAAAGGAGGAAGTACT 695
QY
603 ACAGGGGGAAGCTGTAAAGAAAGAGTCCAGCTGTTGTTAAAGAGGAGGAAGAGTCT 662
Db
696 TCATCAGTTGTCGGAACAGACATCAGATGCTGTGCTAGAAATGCACCTGGCGCGAATG 755
QY
663 TCATGAATTTATTGGATGGATGAAGTGTGCTGCGCAAGAAATGAATGGGAGGGAATG 722
Db
756 CACCAAGCTGTTACCATAGATCTAAACAGAGCTCTCAGAGCACACTCCCCAGTTATT 815
QY
723 CACCAAGCTGTTCAAAATAGATCTAAATCAAACTCTCAGAGCACACTGCCAAATATT 782
Db
816 CCTGAAAGTGGGGAACCCCTTGTGATCAGTGTGAAGGCCATCCATGTGAACATGGATT 875
QY
783 TCTTAAAGTGAAGGGAACCCCTTATGGATAAGTGTGAAGCTGTGATGTGAACCATGATT 842
Db
876 CGGGCTCACCTGGAGCTGGAAGACAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAG 935
QY
843 CGGGCTCACCTGGAAATTAAGAAACAAAGCACTCGAGGAGGCGCACTTTGAGATGAG 902
Db
936 TACCTACTCCACAAACAGAGCATGATTCGGATTCTCTTGGCCCTTTGTGCTCTCCGTTGG 995
QY
903 TACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCAGTGGC 962
Db
996 AAGGAACGACACCGGATATTACACCTGCTTCCTCAAAAGCACCCAGCCAGTCACGGTT 1055
QY
963 AAGAAACGACACCGGATACTACACTGTGTCCTCTTCAAAAGCATCCCAAGTCAATCAGCTTT 1022
Db
1056 GTGACCATCTAGAAAAAGGTTTATAACGCTACCAAGCTCCAGAGAGAGTATCAAAAT 1115
QY
1023 GGTTACCATCTGTAAGAAAGGATTTATAATGCTACCAATTTAAGTGAAGATTATGAAT 1082
Db
1116 TGACCCGCTACGAAAGTCTGCTTCTCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATG 1175
QY
1083 TGACCAATTAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTTACCCACAATCAGATG 1142
Db
1176 CAGTGGATCTCTCTCAGCCCTCATTTCTGTTGAACAGAGAGGCGCTGGAGATGGGTA 1235
QY
1143 TACGTGGACCTCTCTCGAAAAATCAATTCCTGTGAGCAAAAGGGTCTTGATAACGGATA 1202
Db
1236 CAGCATATCAATTTTGGATCATTAAGAACAGCCAGGAGAGTACATATCTCATCAGA 1295
QY
1203 CAGCATATCAAGTTTGGATCATTAAGAACAGCCAGGAGGAATATATTTCCATCAGA 1262
Db
1296 AAATGATGAGCCAGTTTCAACCAAAATGTTACAGCTGAATATAAGAAAGAAACCTCAAGT 1355
QY

1263 AAATGATGATGCCCAATTTTACCAAAATGTTTACGCTGAATATAGAAGGAACCTCAAGT 1322
Db
1356 GCTAGCAATGCTCAGCCAGCCAGCGCTCTCTCTCTGATGGCTACCCGCTACCTC 1415
QY
1323 GCTCGAGAAGTATGGCAAGTCAAGCGCTCTCTCTCTGGAATGGATACCATTTACCATC 1382
Db
1416 TTGGACCTGGAAGAAGTGTTCGGACAAATCTCCAAATTCACGAGGAGAAATCCAGAAAG 1475
QY
1383 TTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAAGAGATCACAGAAG 1442
Db
1476 AGTTTGGAAATAAAAGGCTAAACAGAAAAGTGTTCGCAAGTGGGTGTCGAGCAGTACTCT 1535
QY
1443 AGTCTGGAATAGAAAGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCT 1502
Db
1536 AAATATGAGTGAAGCGCGGAAAGGCTTCTGGTCAAAATGCTGCTGCTACAAATTTCTATGG 1595
QY
1503 AAACATGAGTGAAGCCATAAAAGGTTCTGGTCAAGTGTGTCATACAAATTTCCCTTGG 1562
Db
1596 CACGCTCTCGAAACCATCTTTTAAACTCACAGGCCCTTCCCTTTTCATCCAAGACAA 1655
QY
1563 CACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAA 1622
Db
1656 CATCTCTTCTATGCGACCATTTGGGCTCTGCTCCCTTTCATTTGTTTCTCATFTGTT 1715
QY
1623 CATCTCATTTCTATGCAACAATTTGTTGCTCTCTCTCTCATTTGCTGTTTAAACCTGCT 1682
Db
1716 GATCTGCCACAATAACAAAAGCAATTTAGGTACGAGAGTCAGCTCAGATGATCCAGGT 1775
QY
1683 AATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCACTACAGATGTTACAGGT 1742
Db
1776 GACTGCCCTCTGGATAACGAGTACTTCTACGTTGACTTCAGGAGCTATGAATATGACCT 1835
QY
1743 GACCGCTCTCAGATATGAGTACTTCTACGTTGATTTCAGAGAATATGAATATGATCT 1802
Db
1836 TAAGTGGGAGTTCGCCAGAGAGAACTTAGAGTTTGGGAAGTCTCTGGGCTCTGGCGCTTT 1895
QY
1803 CAAATGGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTT 1862
Db
1896 CGGAGGCTGATGAAGCCAGCCGCTTATGCAATTTAGTAAACGGAGCTCTCAATTCAGGT 1955
QY
1863 TGGAAAGTGTGAAGCAGCACTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1922
Db
1956 GCGGTGAAGATGCTAAAAGAGAAAGTACAGCTGTGAAAAGAAAGCTCTCATGTCCGA 2015
QY
1923 TGCGCTCAAAATGCTCAAGAAAAGCAGACACTCTGAAGAGAGAGCACTCATGTCTAGA 1982
Db
2016 GCTCAAAATGATGACCCACTGGGACACCATGACAAATCTGTAATCTGCTGGGGGATG 2075
QY
1983 ACTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTGTGAACCTGCTGGGGGCTG 2042
Db
2076 CACACTGTGAGGCCAGTGTACTTTGATTTTGAATATTGTTGCTATGCTGACCTCTCA 2135
QY
2043 CACACTGTGAGGCCAGTGTACTTTGATTTTGAATATTGTTGCTATGCTGATCTTCTCA 2102
Db
2136 CTACCTAAGAAGTAAAGAGAGAAGTTTACAGACATGACAGAGATTTTAAAGAAACA 2195
QY
2103 CTATCTAAGAAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGAAACA 2162
Db
2196 TAATTTTCACTTCACTTCTTCCAGGCAATTTCAAAATCCAGCATGCTGTTTCAAG 2255
QY
2163 CAATTTTCACTTCTTCACTTCTTCCAAATTCACATCCAAATTTCCAGCATGCTGTTTCAAG 2222
Db
2256 AGAAGTTTCACTTCTTCACTTCTTCCAGGCTTCCAGGTTCAATGGGAAATTCAAATCA 2315
QY
2223 AGAAGTTTCACTTCTTCACTTCTTCCAGGCTTCCAGGCTTCCAGGAAATTCATTTCA 2282
Db
2316 TTTCTGAAGTGAAGTGAATATCAAAACAGAGAGGCTGGCAAGAAAGAGGAGAGA 2375
QY
2283 CTCTGAAGTGAATTTGAATATGAACCAACCAAAAGGCTG-----GAAGAGAGGAGA 2336
Db
2376 TTTTGAAGTGTGACGTTTGAAGACCTCTTTGCTTTCGTTACCAAGTGGCCAAAGCAT 2435
QY

Search completed: May 25, 2003, 09:18:59
Job time : 8670.83 secs

```
Db 2337 CTTGAATGCTTACATTTGAAGATCTCTTTGCTTTGCATATCAAGTTGCCAAAGGAAT 2396
QY 2436 GGAATTCCTGGAGTTCAAGTGTGTCCACAGAGACCTGGCAGCCAGGAATGCTTGGT 2495
Db 2397 GGAATTCCTGGATTTAAGTGTGTTCACAGAGACCTGGCCGCCAGGAACGTGCTTGT 2456
QY 2496 CACCACCGGGAAGGTGGTGAAGATCTGTGACTTTGGACTGGCCCGAGACATCCTGAGCGA 2555
Db 2457 CACCACCGGGAAGGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGA 2516
QY 2556 CTCAGCTACGTCGACGGGCAACGACAGCGGTGCGGTGAAGTGGATGGCACCCGAGAG 2615
Db 2517 TTCCAACTATGTGTGACGGGCAATGCCCGTGTGCTGTAAATGGATGGCCCGCAAG 2576
QY 2616 CTTATTTGAAGGATCTACAAATCAAGAGTGCCTGTGGCTACGGCATCCTTCTCTG 2675
Db 2577 CTTGTTGAGGCATCTACACCATTAAGATGATGTCTGGTCATATGAATATTACTGTG 2636
QY 2676 GGAGATATTTTCACTGGGTGTGAACCCCTTACCCTGGCATTCCTGCGACGCTAACTTCTA 2735
Db 2637 GGAATCTTCTCACTTGTGTGAATCCTTACCCTGGCATTCGGTGTGATGCTAACTTCTA 2696
QY 2736 TAAACTGATTCAGAGTGAATTTAAATGGAGAGCCATTTCTATGCCACAGAGGATATA 2795
Db 2697 CAAACTGATTCAAATGATTTAAATGGATCGCCATTTTATGCTACAGAAGAAATATA 2756
QY 2796 CTTTGTATGCAATCCTGCGGCTTTTGAAGTGCAGGAGCGGCCATCCTTCCCAACCT 2855
Db 2757 CATTAATGCAATCCTGCTGGCTTTTGAAGTGCAGGAGCGGCCATCCTTCCCTAATTT 2816
QY 2856 GACTTCAATTTTAGGATCTCAGCTGGCAGAGCGCAAGAAG----- 2896
Db 2817 GACTTCGTTTTAGGATCTCAGCTGGCAGATGCAGAAGCGGATGTATCAGATGTGGA 2876
QY 2897 -----CATGTATCAGAATCCATCCATCAAAACAGCGGCCCTCAGCAGAGGC 2952
Db 2877 TGGCCGTGTTTCGGAATGCTCTCACACCTACCAAAACAGCGGCCTTTTCAGCAGAGAT 2936
QY 2953 GGGCTCAGAGC-CCAGTCGCCACAGCGCCAGCT-GAAGATTCACAGAGAAAGATTAGC 3010
Db 2937 GGATTTGGGGTACTCTCTCGCAGGCTCAGTGGAGATTCTGTAGAGGAACAATTTAGT 2996
QY 3011 GAGGAGCCCTTGGACCC-----CGCCACCTAGCAGGCTGTAGACCGCAGAGCCAAAGATTA 3066
Db 2997 TTTAAGGACATTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTA 3056
QY 3067 GCCTCGCCTCT- -GAGGAGCGCCCTACAGCGGTTGCTTCGCTGGAGCTTTTCTCTAGAT 3124
Db 3057 ATTTTCATCACTAAAGAAATCTATTATCAACTGCTCTCACCAGACTTTTCTCTAGAA 3116
QY 3125 GCTGCTGCCATTACTC-----CAAAGTCACTTCTATAAATCAAAACC-TCCTCCGCG 3176
Db 3117 GCTGCTCGGTTTACTCTGTTTCAAGGGGACCTTTGTAAATCAAAATCATCTCTGTAC 3176
QY 3177 ACAGCGGGGAGAGCAATATGAGACTTGTGGTGAGCGCCCTACCCCTGGGGGCTTTC 3236
Db 3177 AAGCAGGAGAGCTGATATGAATCTTATTTGGAGCATGTATCTGCATCCAAGGCCCTCT 3236
QY 3237 CACGAGCTTGAGGGAAGGCCATGTATCTGAAATATAGTATATTTCTGTAATAACGTGAA 3296
Db 3237 CAGGCTGGCTTGAGTGAATTG-TGTACCTGAAGTACAGTATATCTTCTGTAATAACATAA 3295
QY 3297 ACAAAACCAACCCGTTTTTTCCTAAGGGAAGCTAAATATGATTTTTTAAATCTATGTT 3356
Db 3296 ACAAA-----AGCATTTTGTCTAAGGAGAGCTAAATATGATTTTTT-AAAGCTATGTT 3346
QY 3357 TTAATACTATGTAACTTTTTTCATCTATTTAGTGATATATTTTATGATGGAATTAAC 3416
Db 3347 TTAATAATATGTAATAATTTTTTACGCTATTTAGTGATATATTTTATGGTGGGAATAAA 3406
QY 3417 TTTCTACTG 3425
Db 3407 TTTCTACCG 3415
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 04:06:12 ; Search time 695,168 Seconds
(without alignments)
11185.997 Million cell updates/sec

Title: US-09-919-408-1
Perfect score: 3453
Sequence: 1 GCGGCTGGTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	16	AAQ81012
2	3453	100.0	3453	17	AAQ81012
3	3453	100.0	3453	18	AAQ81012
4	3453	100.0	3453	20	AAQ81012
5	3451.4	100.0	3453	14	AAQ81012
6	3451.4	100.0	3453	14	AAQ81012
7	3451.4	100.0	3453	16	AAQ81012
8	3448.2	99.9	3453	13	AAQ81012
9	3446.6	99.8	3453	14	AAQ81012

10	3342.4	96.8	3521	16	AAQ81012	Flk2/flt3 tyrosine
11	2227.2	64.5	3501	16	AAQ81013	Flk2 receptor prot
12	2227.2	64.5	3501	16	AAQ79069	Human flk-2 cDNA.
13	2227.2	64.5	3501	17	AAQ79069	Murine foetal live
14	2227.2	64.5	3501	18	AAQ72117	Human flk-2 recep
15	2227.2	64.5	3501	20	AAQ77515	Human flk-2 cDNA.
16	2225.6	64.5	3501	14	AAQ83503	Human flk-2 cDNA.
17	2225.6	64.5	3501	14	AAQ40915	Human flk-2 cDNA.
18	2224	64.4	3501	14	AAQ35250	Human flk-2 coding
19	2204.8	63.9	3476	16	AAQ91536	Human STK-1 cDNA.
20	2198.2	63.7	3475	16	AAQ08002	Human Flk2/flt3 ty
21	2085	60.4	3120	14	AAQ49756	PTK gene LptK25.
22	2081.8	60.3	3120	16	AAQ03096	Protein tyrosine-k
23	2024.6	58.6	2949	19	AAV39041	Human receptor typ
24	2018.8	58.5	2958	19	AAV39042	Human receptor typ
25	2017.2	58.4	2958	19	AAV39040	Human receptor typ
26	1994.8	57.8	2982	19	AAV39039	Human receptor typ
27	1984.2	57.5	2978	19	AAV39038	Human receptor typ
28	1199.6	34.7	1894	15	AAQ54036	Flk-2ws gene. Mus
29	1119.6	32.4	2247	19	AAV55294	Sequence pMON32390
30	357.4	10.4	3992	23	AAQ79666	DNA encoding novel
31	357.4	10.4	3992	24	AAQ48105	Human macrophage c
32	332.6	9.6	3069	21	AAQ44718	Bovine c-kit bk-1
33	331.4	9.6	2919	20	AAV80687	Hampshire porcine
34	329	9.5	5084	19	AAV20443	Human c-kit Oncoge
35	329	9.5	5084	22	AAQ13426	Human Kit/stem cel
36	329	9.5	5084	24	ABL64113	Breast cancer rela
37	329	9.5	5084	24	ABL64113	Ovary cancer relat
38	325.4	9.4	6390	23	AAQ84936	DNA encoding novel
39	325.4	9.4	6390	23	AAQ84936	DNA encoding novel
40	325	9.4	5098	22	AAQ13425	Murine Kit/stem ce
41	286.4	8.3	2463	20	AAQ89336	Platelet-derived g
42	286.4	8.3	4054	11	AAQ06869	Sequence encodes p
43	286.4	8.3	4544	10	AAQ90355	cDNA encoding plat
44	286.4	8.3	4544	17	AAQ34552	Platelet-derived g
45	286.4	8.3	5602	23	AAQ84940	DNA encoding novel

ALIGNMENTS

RESULT 1	
AAQ81012	
ID	AAQ81012 standard; cDNA; 3453 BP.
XX	AAQ81012;
AC	AAQ81012;
XX	
DT	16-AUG-1995 (first entry)
XX	
DE	Flk2 receptor protein-tyrosine-kinase cDNA.
XX	
KW	Mouse Flk2; receptor protein-tyrosine-kinase; primitive
KW	hematopoietic cell; fetal liver kinase; ds.
OS	Mus musculus.
XX	
FH	Key
CDS	Location/Qualifiers
FT	31..3009
FT	/*tag= a
FT	/product= Flk2 receptor protein-tyrosine-kinase
FT	31..111
FT	/*tag= b
FT	112..3006
FT	/*tag= c
XX	
PN	WO95005554-A.
XX	
PD	05-JAN-1995.
XX	
PF	17-JUN-1994;
XX	94WO-US06944.
PR	18-JUN-1993;
PR	93US-0080244.
PR	21-JUN-1993;
PR	93US-0081508.

Db	2821	TTTGACTCAAGGAAGCGGCATCTCTCCCAACTGACTTCATTTTATGAGTGTACGCTG	2888
Qy	2881	GCAGAGCGAAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGCCCT	2940
Db	2881	GCAGAGCGAAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGCCCT	2940
Qy	2941	CAGCAGAGAGCGGGCTCAGAGCCCATGTGCGCACAGCGCAGGAGATTCACAGAA	3000
Db	2941	CAGCAGAGAGCGGGCTCAGAGCCCATGTGCGCACAGCGCAGGAGATTCACAGAA	3000
Qy	3001	AGAAAGTTAGCGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Db	3001	AGAAAGTTAGCGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Qy	3061	AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT	3120
Db	3061	AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT	3120
Qy	3121	AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG	3180
Db	3121	AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG	3180
Qy	3181	CGCGGAGAGCCAATAATGAGACTTGTGGTGAGCCCGCTACCTGGGGGCTTTCACG	3240
Db	3181	CGCGGAGAGCCAATAATGAGACTTGTGGTGAGCCCGCTACCTGGGGGCTTTCACG	3240
Qy	3241	AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTGCTAAATACGTGAAACAA	3300
Db	3241	AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTGCTAAATACGTGAAACAA	3300
Qy	3301	ACCAAAACCCGTTTTTCTGAAGGAAAGCTAAATATATGATTTTAAAAATCTATGTTTAA	3360
Db	3301	ACCAAAACCCGTTTTTCTGAAGGAAAGCTAAATATATGATTTTAAAAATCTATGTTTAA	3360
Qy	3361	AATGCTATGTAACCTTTTTCATCTATTTAGTCATATATTTATGATGGAATAAATCTTC	3420
Db	3361	AATGCTATGTAACCTTTTTCATCTATTTAGTCATATATTTATGATGGAATAAATCTTC	3420
Qy	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453	
Db	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453	
RESULT 2			
AAT38733			
ID	AAT38733	standard; cdna; 3453 BP.	
XX	AC	AAT38733;	
XX	DT	11-DEC-1996 (first entry)	
XX	DE	Human foetal liver kinase 2 cDNA.	
KW		Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;	
KW		monoclonal; antibody; extracellular domain; receptor assay;	
KW		haematopoietic stem cell; ligand; stimulation; proliferation;	
KW		differentiation; treatment; anaemia; bone marrow damage;	
KW		cancer chemotherapy; radiation; ds.	
XX	OS	Homo sapiens.	
XX	FH	Key	Location/Qualifiers
FT	CDS	31..3009	
FT		/*tag= a	
FT	sig_peptide	31..111	
FT		/*tag= b	
FT	mat_peptide	112..3006	
FT		/*tag= c	
XX			
PN		US5548065-A.	
XX	PD	20-AUG-1996.	
XX			

PF	02-APR-1991;	91US-0679666.
XX		
PR	19-NOV-1992;	92US-0977451.
PR	02-APR-1991;	91US-0679666.
PR	28-JUN-1991;	91US-0728913.
PR	15-NOV-1991;	91US-0793065.
PR	24-DEC-1991;	91US-0813593.
PR	26-JUN-1992;	92US-0906397.
PR	12-NOV-1992;	92US-0975049.
PR	30-APR-1993;	93US-0055269.
PR	31-OCT-1994;	94US-0252517.
XX	(UYPR-) UNIV PRINCETON.	
XX	Lemischka IR;	
PI		
XX	WPI; 1996-392678/39.	
DR	P-PSDB; AAR97418.	
XX		
PT	Anti-fœtal liver kinase 2 (flk-2) antibodies - useful in assays,	
PT	for isolating haematopoietic stem cells expressing receptor and for	
PT	obtaining ligands	
XX		
PS	Disclosure; Columns 27-34; 50pp; English.	
XX		
CC	The present sequence encodes human foetal liver kinase 2 (flk-2), a	
CC	protein tyrosine kinase. Isolated antibodies, pref. monoclonal,	
CC	raised against the extracellular portion of flk-2 can be used to	
CC	assay for flk receptors on the surface of primitive haematopoietic	
CC	stem cells, and to isolate positive cells. The antibodies can also	
CC	be used as, or to obtain ligands, which stimulate the proliferation	
CC	and/or differentiation of stem cells. The ligands can be used, e.g.	
CC	for treating anaemia, or bone marrow damage resulting from cancer	
CC	chemotherapy, or radiation.	
XX		
SQ	Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;	
	Query Match 100.0%; Score 3453; DB 17; Length 3453;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GCGGCCTGGGTACCGCGGTCGCGAGGCCATGCGGGCGTTGGCGCAGCGCAGCGCGG 60	
Dd	1 GCGGCCTGGGTACCGCGGTCGCGAGGCCATGCGGGCGTTGGCGCAGCGCAGCGCGG 60	
Qy	61 CGGCTGCTGCTGCTTGTGTTTCTCAGTAATGATTCTTGAGACCGTTACAACCAAGAC 120	
Dd	61 CGGCTGCTGCTGCTTGTGTTTCTCAGTAATGATTCTTGAGACCGTTACAACCAAGAC 120	
Qy	121 CTGCTGTGATCAAËGTGTTTTAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG 180	
Dd	121 CTGCTGTGATCAAËGTGTTTTAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG 180	
Qy	181 CCATCATCGTACCAGATGGTCGAGGATPCCCAGAACCTCCAGTAGTACCCGAGCGGC 240	
Dd	181 CCATCATCGTACCAGATGGTCGAGGATPCCCAGAACCTCCAGTAGTACCCGAGCGGC 240	
Qy	241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGCTGGGTCCATC 300	
Dd	241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGCTGGGTCCATC 300	
Qy	301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGACCTTTCCTGCCTCTGGGTCTTTAAGCAC 360	
Dd	301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGACCTTTCCTGCCTCTGGGTCTTTAAGCAC 360	
Qy	361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTACAAAACAGAGGAATGTTTCCATGGCC 420	
Dd	361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTACAAAACAGAGGAATGTTTCCATGGCC 420	
Qy	421 ATCTTTGAACGTTGACAGACCCAGCGAGAGAATACCTACTTCATATTCAGAGCGAAGCG 480	
Dd	421 ATCTTTGAACGTTGACAGACCCAGCGAGAGAATACCTACTTCATATTCAGAGCGAAGCG 480	

Key	Location/Qualifiers	Seq
Key CDS	31..3009	241 CAGAGTGAAGGACGGTATATGAAGCCGCCACCCACCTGGAGGTGGCCGAGTCTGGGTCCATC 300
FT	/*tag= a	301 ACCCTGCAAGTGCAGCTCGCCACCCACAGGGACACCTTTCTGCTCTCTGGGTCTTTAAAGCAC 360
FT	/product= Murine_flk-2	301 ACCCTGCAAGTGCAGCTCGCCACCCACAGGGACACCTTTCTGCTCTCTGGGTCTTTAAGCAC 360
FT	31..111	301 ACCCTGCAAGTGCAGCTCGCCACCCACAGGGACACCTTTCTGCTCTCTGGGTCTTTAAGCAC 360
FT	/*tag= b	361 AGCTCCCHGGGCTGCCAGCGCACTTTGATTTTACAAAACAGAGAGAAATCGTTTTCATGGCC 420
FT	112..3006	361 AGCTCCCTGGGCTGCCAGCGCACTTTGATTTTACAAAACAGAGAGAAATCGTTTTCATGGCC 420
FT	/*tag= c	361 AGCTCCCTGGGCTGCCAGCGCACTTTGATTTTACAAAACAGAGAGAAATCGTTTTCATGGCC 420
PN	US5621090-A.	421 ATCTTGAACTGACAGAGACCCAGCGAGGAGAATACCTACTCCATATTCACAGAGGAACGC 480
PN	15-APR-1997.	421 ATCTTGAACTGACAGAGACCCAGCGAGGAGAATACCTACTCCATATTCACAGAGGAACGC 480
PP	02-APR-1991;	481 GCCAACTACACAGTACTGTTCCACAGTGAATTAAGAGAGATACACAGCTGCTATGCTAAGG 540
PP	26-JUN-1992;	481 GCCAACTACACAGTACTGTTCCACAGTGAATTAAGAGAGATACACAGCTGCTATGCTAAGG 540
PR	91US-0679666.	481 GCCAACTACACAGTACTGTTCCACAGTGAATTAAGAGAGATACACAGCTGCTATGCTAAGG 540
PR	02-APR-1991;	541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGGTGTT 600
PR	28-JUN-1991;	541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGGTGTT 600
PR	15-NOV-1991;	541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGGTGTT 600
PR	91US-0793065.	541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGGTGTT 600
PR	24-DEC-1991;	541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGGTGTT 600
XX	(UYPR-) UNIV PRINCETON.	601 CCGGAGCCCACTGTGGAGTGGTCTCTGCAAGTCCACAGGGAAGAGCTGTAAAGAAGAA 660
XX	PA	601 CCGGAGCCCACTGTGGAGTGGTCTCTGCAAGTCCACAGGGAAGAGCTGTAAAGAAGAA 660
XX	Lemischka IR;	601 CCGGAGCCCACTGTGGAGTGGTCTCTGCAAGTCCACAGGGAAGAGCTGTAAAGAAGAA 660
PI	WPI; 1997-235228/21.	661 GGCCTGCTGTTGTCAGAAAGGAGGAAAGTACTTTCATGAGTTGTTTCGGAACACAGATC 720
DR	P-PSDB; AAW19874.	661 GGCCTGCTGTTGTCAGAAAGGAGGAAAGTACTTTCATGAGTTGTTTCGGAACACAGATC 720
XX	Protein containing the extracellular domain of human flk-2 - used for identification of primitive haematopoietic cell proliferation and differentiation stimulatory ligands, e.g. for treating anaemia	721 AGATGCTGTGCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTTCCACCATAGATCTA 780
PPT	Disclosure; Fig 1A; 55pp; English.	721 AGATGCTGTGCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTTCCACCATAGATCTA 780
PPT	This sequence encodes the murine fetal liver kinase 2 (flk2). flk-2 is a receptor protein tyrosine kinase (pTK) and is important in transducing putative self-renewal signals from the environment. flk-2 is expressed in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow, and it is thought that flk-2 is expressed in the entire primitive portion of the haematopoietic hierarchy. The invention concerns a recombinant nucleic acid, preferably mRNA, which encodes a protein containing only the extracellular domain of human flk-2 and lacking the flk-2 intra-cellular catalytic domain. The resultant protein represents a soluble form of flk-2 which is used to isolate specific ligands for flk-2. These ligands can be used to stimulate proliferation and/or differentiation of mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for treatment of macrocytic or aplastic anaemia or bone marrow damage caused by cancer treatment or radiation.	781 AACAGGCTGAGGAGGAGGAGTCTTTCAGATGAGTACTTCCACCAACAGAGACCATG 960
XX	Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;	781 AACAGGCTGAGGAGGAGGAGTCTTTCAGATGAGTACTTCCACCAACAGAGACCATG 960
XX	Query Match 100.0%; Score 3453; DB 18; Length 3453;	901 AAAGCCCTGGAGGAGGAGGAGTCTTTCAGATGAGTACTTCCACCAACAGAGACCATG 960
XX	Best Local Similarity 100.0%; Pred. No. 0;	901 AAAGCCCTGGAGGAGGAGGAGTCTTTCAGATGAGTACTTCCACCAACAGAGACCATG 960
XX	Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	961 ATTCGGATCTCTTGGCCCTTTGTTGTTCTTCCTGGGAGGAGACACCGGATATTACACC 1020
XX		961 ATTCGGATCTCTTGGCCCTTTGTTGTTCTTCCTGGGAGGAGACACCGGATATTACACC 1020
XX		1021 TGCTCTTCTCAAGACACCCAGCCAGCTGAGCGTGGTGGATCTCTCAAGAAAGGGTTT 1080
XX		1021 TGCTCTTCTCAAGACACCCAGCCAGCTGAGCGTGGTGGATCTCTCAAGAAAGGGTTT 1080
QY	1 GCGGCTGGCTACCGCGCTCGGAGGCCATCGGCGGTTGGCGCAGCGACGACCGG 60	1081 ATAAACGCTTACAGCTCGCAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1140
QY	1 GCGGCTGGCTACCGCGCTCGGAGGCCATCGGCGGTTGGCGCAGCGACGACCGG 60	1081 ATAAACGCTTACAGCTCGCAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1140
QY	61 CGGCTGCTGCTGTTGTTTTCAGTAAATGATTTCTGAGACCGTTTACAAACCAAGAC 120	1141 TCAGTCAGTTTAAAGCGTACCCAGCATCCGATCGATGCTCTCTCAAGCCCTCA 1200
QY	61 CGGCTGCTGCTGTTGTTTTCAGTAAATGATTTCTGAGACCGTTTACAAACCAAGAC 120	1141 TCAGTCAGTTTAAAGCGTACCCAGCATCCGATCGATGCTCTCTCAAGCCCTCA 1200
QY	121 CTGCGCTGATCAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGGAAG 180	1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY	121 CTGCGCTGATCAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGGAAG 180	1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY	181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTTCCAGTGTACCCGAGCGC 240	1261 AAGAACACAGCCAGGAGAGTACATATTTCTATCGAAGAAATGATGACGCCAGTTCACCAA 1320
QY	181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTTCCAGTGTACCCGAGCGC 240	1261 AAGAACACAGCCAGGAGAGTACATATTTCTATCGAAGAAATGATGACGCCAGTTCACCAA 1320
QY	241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300	1321 ATGTTTACAGCTGATATTAAGAAAGAACCTCAAGTGTAGCAAAATGATGACGCCAGTTCACCAA 1380
QY		1321 ATGTTTACAGCTGATATTAAGAAAGAACCTCAAGTGTAGCAAAATGATGACGCCAGTTCACCAA 1380

```
QY 1381 GCGTCTCTGTTCTCTGATGGCTTACCGCTTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440
Db 1381 GCGTCTCTGTTCTCTGATGGCTTACCGCTTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440
QY 1441 AAATCTCCAAATTGACGAGGAAATCCAGAGGAGTTTGAATATAAAGGCTAACAGA 1500
Db 1441 AAATCTCCAAATTGACGAGGAAATCCAGAGGAGTTTGAATATAAAGGCTAACAGA 1500
QY 1501 AAAGTGTTCGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560
Db 1501 AAAGTGTTCGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560
QY 1561 CTTCCTGGTCAAAATGCTGTGCGTACAAATCTATGGGCACGCTCTGGGAAACCATCTTTT 1620
Db 1561 CTTCCTGGTCAAAATGCTGTGCGTACAAATCTATGGGCACGCTCTGGGAAACCATCTTTT 1620
QY 1621 AACTCACAGGCCCCCTCCCTTTCATCCAAAGACAATCTCTTCTATGCGACCATTTGGG 1680
Db 1621 AACTCACAGGCCCCCTCCCTTTCATCCAAAGACAATCTCTTCTATGCGACCATTTGGG 1680
QY 1681 CTCGTCTCCCTTCAATTTGTTCTCATTTGTTGATCTGATCTGCCCAAAATACAAAAGCAA 1740
Db 1681 CTCGTCTCCCTTCAATTTGTTCTCATTTGTTGATCTGATCTGCCCAAAATACAAAAGCAA 1740
QY 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGGCCCTGGATACAGAGTAC 1800
Db 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGGCCCTGGATACAGAGTAC 1800
QY 1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
QY 1861 TTAGAGTTTGGGAAGTCTGGGGTCTGGCGTCTTCGGAGGGTGATGAACGCCAGGCC 1920
Db 1861 TTAGAGTTTGGGAAGTCTGGGGTCTGGCGTCTTCGGAGGGTGATGAACGCCAGGCC 1920
QY 1921 TATGTCATTAGTAAACGGGAGTCTCAATTACAGGTGGCGTCAAGATGCTTAAAGAGRAA 1980
Db 1921 TATGTCATTAGTAAACGGGAGTCTCAATTACAGGTGGCGTCAAGATGCTTAAAGAGRAA 1980
QY 1981 GGTACAGCTGTGAAAAAGAACTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GGTACAGCTGTGAAAAAGAACTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
QY 2101 ATTTTGTGAATATTGTTGCTATGCTGACCTCCCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
Db 2101 ATTTTGTGAATATTGTTGCTATGCTGACCTCCCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
QY 2161 TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220
Db 2161 TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220
QY 2221 CAGGCACATTTAAATTCAGCATGCTGTTTCAAGAGAGTTTCAATTTACACCCCGCCCTTG 2280
Db 2221 CAGGCACATTTAAATTCAGCATGCTGTTTCAAGAGAGTTTCAATTTACACCCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGGTTCATCGGAATTCATCTGAGATGAGATTCGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCATCGGAATTCATCTGAGATGAGATTCGAATATGAA 2340
QY 2341 AACCAAGAGAGGCTGGCAGAAAGAGGAGGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400
Db 2341 AACCAAGAGAGGCTGGCAGAAAGAGGAGGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTTGGAGTCAAGTCGTG 2460
Db 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTTGGAGTCAAGTCGTG 2460
```

RESULT 4
AAX77514
ID AAX77514 standard; cDNA; 3453 BP.
XX

```
2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCACCGGAGGTGGTGAAGATC 2520
2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCACCGGAGGTGGTGAAGATC 2520
2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2580
2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2580
2581 GCACGCTCCCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640
2581 GCACGCTCCCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640
2641 AAGAGTGACGCTGTCCTACGGCATCCTCTCTGGGAGATATTTTTCACCTGGGTGTGAAC 2700
2641 AAGAGTGACGCTGTCCTACGGCATCCTCTCTGGGAGATATTTTTCACCTGGGTGTGAAC 2700
2701 CCTTACCTGGCAATCCTGTCGACGCTAACTTCTATAAAGTGAATTTTAA 2760
2701 CCTTACCTGGCAATCCTGTCGACGCTAACTTCTATAAAGTGAATTTTAA 2760
2761 ATGGAGCAGCCATTTCTATGCCACAGAGGATATCTTCTGTAATGCAATCCTGCTGGCT 2820
2761 ATGGAGCAGCCATTTCTATGCCACAGAGGATATCTTCTGTAATGCAATCCTGCTGGCT 2820
2821 TTTGACTCAAGGAAGCGGCATCCTTCCCAACCTGACTTCTATTTTAGGATGTCACTG 2880
2821 TTTGACTCAAGGAAGCGGCATCCTTCCCAACCTGACTTCTATTTTAGGATGTCACTG 2880
2881 GCAGAGGCAAGAACATGTATCAGAACATCCATCCATCAGAACATCCATCCATCAGAAC 2940
2881 GCAGAGGCAAGAACATGTATCAGAACATCCATCCATCAGAACATCCATCCATCAGAAC 2940
2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGTGAAGATTCACAGAGAA 3000
2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGTGAAGATTCACAGAGAA 3000
3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCCAGTCCAGCGCTAGACAGCGTGAAGATTC 3060
3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCCAGTCCAGCGCTAGACAGCGTGAAGATTC 3060
3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGCTTCTGCTGCTGAGCTTTTCTCT 3120
3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGCTTCTGCTGCTGAGCTTTTCTCT 3120
3121 AGATGCTGTGCTGCTTACTCCAAAGTACTTCTATAAATCAAACTCTCTCGCACAG 3180
3121 AGATGCTGTGCTGCTTACTCCAAAGTACTTCTATAAATCAAACTCTCTCGCACAG 3180
3181 GCGGGAGAGCAATAATGAGACTTGTGGTGAGCCGCCCTACCTGGGGCTTTTCCACG 3240
3181 GCGGGAGAGCAATAATGAGACTTGTGGTGAGCCGCCCTACCTGGGGCTTTTCCACG 3240
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
3301 ACCAAACCCGTTTTTTCGCTAAGGGAAGCTTAAATATGATGATTTTAAATCTATGTTTAA 3360
3301 ACCAAACCCGTTTTTTCGCTAAGGGAAGCTTAAATATGATGATTTTAAATCTATGTTTAA 3360
3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGATGGAATTAACCTTC 3420
3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGATGGAATTAACCTTC 3420
3421 TACTGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3453
3421 TACTGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3453
```

AC AAX77514;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE Murine flk-2 cDNA.
 XX
 KW Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
 XX monoclinal; polyclonal; antibody; tyrosine kinase; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 31..3009
 FT /*tag= a
 FT /product= "flk-2"
 XX
 PN US5912133-A.
 XX
 PD 15-JUN-1999.
 XX
 PF 10-FEB-1998; 98US-0021324.
 XX
 PR 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 DR WPI; 1999-357194/30.
 DR P-PSDB; AAY08616.
 XX
 PT Isolating hematopoietic cells expressing fetal liver kinase 1
 PT receptors
 XX
 PS Disclosure; Fig 1a; 59pp; English.
 XX
 CC This invention describes a novel method of isolating cells expressing
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
 CC binding the cells to a polyclonal or monoclonal antibody specific to
 CC the FLK-1 receptor and isolating the cells that have bound to the
 CC antibody. The method can be used to isolate hematopoietic stem cells in
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
 CC the invention belong to the receptor protein family. This sequence
 CC encodes the murine flk-2 protein which is used in the method of the
 CC invention.
 XX
 SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;
 Query Match 100.0%; Score 3453; DB 20; Length 3453;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCGGCTGGCTACCGCGCGCTCCGGAGGCATCGGGCGTTGGCGCAGCGCAGCGCGG 60
 DB 1 GCGGCTGGCTACCGCGCGCTCCGGAGGCATCGGGCGTTGGCGCAGCGCAGCGCGG 60
 QY 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTAATGATCTTGAGACCGGTTACAAACCAAGAC 120
 DB 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTAATGATCTTGAGACCGGTTACAAACCAAGAC 120
 QY 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAG 180
 DB 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAG 180

181 CCATCATCGTACCGAATGGTGGAGGATCCCAAGAACCTCCAGTGTACCCCGAGGCGC 240
 181 CCATCATCGTACCGAATGGTGGAGGATCCCAAGAACCTCCAGTGTACCCCGAGGCGC 240
 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGGACCTTCCCTGCTCTGCTCTTAAAGCAC 360
 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGGACCTTCCCTGCTCTGCTCTTAAAGCAC 360
 361 AGCTCCCTGGGCTGCCAGCGCAGCTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC 420
 361 AGCTCCCTGGGCTGCCAGCGCAGCTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC 420
 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCCATATTCAGAGGGAAGC 480
 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCCATATTCAGAGGGAAGC 480
 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAGAGATACACAGCTGTATGTCTAAGG 540
 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAGAGATACACAGCTGTATGTCTAAGG 540
 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCAGTCTCTGCATCTCCGAGGGTGT 600
 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCAGTCTCTGCATCTCCGAGGGTGT 600
 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 661 GGCCTGCTGTTGTGAGAAAGGAGAAAGTACTTCTATGATGTTGTTGCGGAACAGACATC 720
 661 GGCCTGCTGTTGTGAGAAAGGAGAAAGTACTTCTATGATGTTGTTGCGGAACAGACATC 720
 721 AGATGCTGTCTAGAAATGCAGTGGCGCGGAATGCACCAAGCTGTTTCAACCATAGATCTA 780
 721 AGATGCTGTCTAGAAATGCAGTGGCGCGGAATGCACCAAGCTGTTTCAACCATAGATCTA 780
 781 AACAGGCTCCTCAGAGCACACTGCCCAAGTATTCTTGAAGTGGGGGAAACCTTGTGG 840
 781 AACAGGCTCCTCAGAGCACACTGCCCAAGTATTCTTGAAGTGGGGGAAACCTTGTGG 840
 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACTGGGAGCTGGAAGAC 900
 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACTGGGAGCTGGAAGAC 900
 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960
 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960
 961 ATTCGGATCTCTTTGGCTTTGTTGCTTCCGTGGGAAGGAACAGACACCGGATATTACACC 1020
 961 ATTCGGATCTCTTTGGCTTTGTTGCTTCCGTGGGAAGGAACAGACACCGGATATTACACC 1020
 1021 TGCTCTTCTTCARAGCACCCCGAGCCAGTGGTGGTGCACCTCTAGAAAAGGGTGT 1080
 1021 TGCTCTTCTTCARAGCACCCCGAGCCAGTGGTGGTGCACCTCTAGAAAAGGGTGT 1080
 1081 ATAAACCTACAGCTCGCAAGAAGAGTATGAATGACCGTACGAAAGTCTTCTCAAGCTTCA 1140
 1081 ATAAACCTACAGCTCGCAAGAAGAGTATGAATGACCGTACGAAAGTCTTCTCAAGCTTCA 1140
 1141 TCAGTCAAGTCTTAAAGCGTACCCACGAATCCGATGCGATGCGATGCGATGCGATGCGAT 1200
 1141 TCAGTCAAGTCTTAAAGCGTACCCACGAATCCGATGCGATGCGATGCGATGCGATGCGAT 1200
 1201 TTTCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
 1201 TTTCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
 1261 AAGNACAGCCAGGAGAGTACATATTCTATGCAGAAATGATGACGCCAGTTCACCAAA 1320

QY 1201 TTTCTTGTGTGAACAGAGAGGCTGAGAGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260
DB 1201 TTTCTTGTGTGAACAGAGAGGCTGAGAGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY 1261 AAGAACAGCCAGGAGAGTACATATTTATCGAGAAATATGATGACGCCAGTTTACCCAAA 1320
DB 1261 AAGAACAGCCAGGAGAGTACATATTTATCGAGAAATATGATGACGCCAGTTTACCCAAA 1320
QY 1321 ATGTTACGCTGAAATATAAGAAAGAAACCTCAAGTGTAGCAATATGCTCAGCCAGCCAG 1380
DB 1321 ATGTTACGCTGAAATATAAGAAAGAAACCTCAAGTGTAGCAATATGCTCAGCCAGCCAG 1380
QY 1381 GCGTCTGTTCCTCTGATGGTACCCGCTACCCCTTGGACCTGGAAGAGTTTTCGGAC 1440
DB 1381 GCGTCTGTTCCTCTGATGGTACCCGCTACCCCTTGGACCTGGAAGAGTTTTCGGAC 1440
QY 1441 AAATCTCCCAATTCACGGAGGAATCCAGAGGAGTTTGGATAAAGGCTTAACAGA 1500
DB 1441 AAATCTCCCAATTCACGGAGGAATCCAGAGGAGTTTGGATAAAGGCTTAACAGA 1500
QY 1501 AAAGTGTTCGCCAGTGGTCTGAGCAGTACTCTAAATATAGTGAAGCGGGGAAAGGG 1560
DB 1501 AAAGTGTTCGCCAGTGGTCTGAGCAGTACTCTAAATATAGTGAAGCGGGGAAAGGG 1560
QY 1561 CTCTGTCAAAATCTGTGCTACAAATTTCTATGGSCAGTCTTGGCAACCATCTTTTA 1620
DB 1561 CTCTGTCAAAATCTGTGCTACAAATTTCTATGGSCAGTCTTGGCAACCATCTTTTA 1620
QY 1621 AACTCAGCAGCCCTTCCCTTTCATCCAGAACACATCTCTTATGGACACATTTGGG 1680
DB 1621 AACTCAGCAGCAGTACGTCAGATGATCAGGTGACTGGCCCTCGATACAGGATAC 1800
QY 1741 TTTAGTACGAGAGTACGTCAGATGATCAGGTGACTGGCCCTCGATACAGGATAC 1860
DB 1741 TTTAGTACGAGAGTACGTCAGATGATCAGGTGACTGGCCCTCGATACAGGATAC 1860
QY 1801 TTCTACGTTGACTTACAGGAGTGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAC 1860
DB 1801 TTCTACGTTGACTTACAGGAGTGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAC 1860
QY 1861 TTAGAGTTTGGGAAGTCTCGGGTCTGGGCTTTGGGAGGGTGAAGCCACAGGCC 1920
DB 1861 TTAGAGTTTGGGAAGTCTCGGGTCTGGGCTTTGGGAGGGTGAAGCCACAGGCC 1920
QY 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA 1980
DB 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA 1980
QY 1981 GCTGACAGTGTGAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCAGTGGGA 2040
DB 1981 GCTGACAGTGTGAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCAGTGGGA 2040
QY 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGACACATGTCAGGCCAGTGTACTTG 2100
DB 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGACACATGTCAGGCCAGTGTACTTG 2100
QY 2101 ATTTTGAATATTTGCTATGTTGACCTCCTCAACTACCTAAGAAAGTAAAGAGAGAG 2160
DB 2101 ATTTTGAATATTTGCTATGTTGACCTCCTCAACTACCTAAGAAAGTAAAGAGAGAG 2160
QY 2161 TTTTACAGGACATGGACAGAGATTTTAAAGAACATATTTTCAAGTTCACCTTCTTC 2220
DB 2161 TTTTACAGGACATGGACAGAGATTTTAAAGAACATATTTTCAAGTTCACCTTCTTC 2220
QY 2221 CAGGCATTCATTAATTCAGATGCTGGTTCAGAGAGAGTTTCAAGTTCACCTTCTTC 2280
DB 2221 CAGGCATTCATTAATTCAGATGCTGGTTCAGAGAGAGTTTCAAGTTCACCTTCTTC 2280
QY 2281 GATCAGTCTCAGGGTTCAATGGGAAATTCATTTCAATTCAGATGAGATTGAATATGAA 2340

DB 2281 GATCAGTCTCAGGGTTCAATGGGAAATTCATTTCAATTCAGATGAGATTGAATATGAA 2340
QY 2341 ACCCAGAGAGGCTGGCAGAAAGAGGAGGAGATTTGAACGTGCTGAGTTTGAGAC 2400
DB 2341 ACCCAGAGAGGCTGGCAGAAAGAGGAGGAGATTTGAACGTGCTGAGTTTGAGAC 2400
QY 2401 CTCCTTTGCTTTGCGTACCAAGTGCCAAAGGCATGGAATTCCTGGAGTTCAGTCTGT 2460
DB 2401 CTCCTTTGCTTTGCGTACCAAGTGCCAAAGGCATGGAATTCCTGGAGTTCAGTCTGT 2460
QY 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCAGGAGGTGTTGAGATC 2520
DB 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCAGGAGGTGTTGAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCGAGACATCCTGAGCGACTCCAGCTACCTGTCAGGGGCAAC 2580
DB 2521 TGTGACTTTGGACTGGCCGAGACATCCTGAGCGACTCCAGCTACCTGTCAGGGGCAAC 2580
QY 2581 GCACGGCTGCCGCTGAAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACAAATC 2640
DB 2581 GCACGGCTGCCGCTGAAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACAAATC 2640
QY 2641 AAGAGTACGCTGTGTCCTACGGCATCTCTTCTCTGGGAGATATTTTCACTGGGTGGAAC 2700
DB 2641 AAGAGTACGCTGTGTCCTACGGCATCTCTTCTCTGGGAGATATTTTCACTGGGTGGAAC 2700
QY 2701 CCTTACCTGGCATCTCTGTCAGCGCTAACTTCTATAACTGATTGAGAGTGAATTTAA 2760
DB 2701 CCTTACCTGGCATCTCTGTCAGCGCTAACTTCTATAACTGATTGAGAGTGAATTTAA 2760
QY 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
DB 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCCTCTGGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880
DB 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880
QY 2881 GCAGAGCAGAGAGACATGATCAGAACATCCATCCATCTACCAAAAGCGGCGCCCT 2940
DB 2881 GCAGAGCAGAGAGACATGATCAGAACATCCATCCATCTACCAAAAGCGGCGCCCT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCAGGTGAAGATTCACAGAGAA 3000
DB 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCAGGTGAAGATTCACAGAGAA 3000
QY 3001 AGAAGTTAGCAGAGAGGCTTGGACCCGCCACCTTAGCAGGCTGTAGACCCGAGAGCCA 3060
DB 3001 AGAAGTTAGCAGAGAGGCTTGGACCCGCCACCTTAGCAGGCTGTAGACCCGAGAGCCA 3060
QY 3061 AGATTAGCCTCGCCTCTGAGGAGCGCCCTACAGCGCTTGGTTCCTGGACTTTTCTCT 3120
DB 3061 AGATTAGCCTCGCCTCTGAGGAGCGCCCTACAGCGCTTGGTTCCTGGACTTTTCTCT 3120
QY 3121 AGATCTGTCTGCCATTACTCCAAAGTGAATCTATAAATCAAACTCTCTCTCGCACAG 3180
DB 3121 AGATCTGTCTGCCATTACTCCAAAGTGAATCTATAAATCAAACTCTCTCTCGCACAG 3180
QY 3181 GCGGAGAGCCCAATTAATGAGACTTGTGTTGAGCCGCCCTACCTGGGGGCGCTTCCACG 3240
DB 3181 GCGGAGAGCCCAATTAATGAGACTTGTGTTGAGCCGCCCTACCTGGGGGCGCTTCCACG 3240
QY 3241 AGCTTGGGGGAAAGCCATGATCTGAATATAGTATATTTCTGTAAATACGTGAACAA 3300
DB 3241 AGCTTGGGGGAAAGCCATGATCTGAATATAGTATATTTCTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCGTTTGTGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
DB 3301 ACCAAACCGTTTGTGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
QY 3361 AATACATGTAATTTTCTATCTATTTAGTATATTTTATGATGGAATAAATCTTC 3420
DB 3361 AATACATGTAATTTTCTATCTATTTAGTATATTTTATGATGGAATAAATCTTC

Db 3361 AATACTATGTAACTTTTTCATCTATTATTAGTGATATATTTTATGGATGGAATAAACTTTC 3420
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
 Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 6

AAQ35249
 ID AAQ35249 standard; cDNA; 3453 BP.

XX AAQ35249;

XX 25-JUN-1993 (first entry)

XX Murine flk-2 coding sequence.

XX Murine; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;
 KW hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-2;
 KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;
 KW multipotential; T-lymphoid; lineage; ss.

XX Mus musculus.

XX Key Location/Qualifiers

PH 31..3009

FT *tag= a

FT 31..111

FT /*tag= b

FT /*note= "Hydrophobic leader"

XX WO9300349-A.

XX 07-JAN-1993.

XX 26-JUN-1992; 92WO-US05401.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 02-APR-1992; 92WO-US02750.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-036323/04.

XX P-PSDB; AAR31375.

XX Nucleic acid encoding receptor protein tyrosine kinase - allows
 PT development of ligands to stimulate proliferation and/or
 PT differentiation of mammalian haematopoietic stem cells

XX Claim 5; Fig 1a; 78pp; English.

XX This sequence encodes a murine receptor protein tyrosine kinase which
 CC belongs to a new functional class of protein tyrosine kinases (pTKs).
 CC pTKs in this class are expressed in primitive mammalian hematopoietic
 CC (pHC) cells but not in mature hematopoietic cells (mHC). The protein
 CC encoded by this sequence is an example of a receptor pTK and is called
 CC fetal liver kinase 2 (flk-2). flk-2 is expressed in fetal liver,
 CC spleen and thymus, and adult brain and marrow. Expression of flk-2
 CC mRNA occurs in the most primitive thymocyte subset, which is believed
 CC to be uncommitted. Therefore, thymocytes expressing flk-2 may be
 CC multipotential. flk-2 is the first receptor tyrosine kinase known to
 CC be expressed in the T-lymphoid lineage.

XX Sequence 3453 BP; 947 A; 821 C; 876 G; 809 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 3451.4; DB 14; Length 3453;

Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCGCTGGCTACCGCGCGCTCCGAGGCGCATCGGGGCTTGGCGAGCGACGACCGG 60
 Db 1 CGGCGCTGGCTACCGCGCGCTCCGAGGCGCATCGGGGCTTGGCGAGCGACGACCGG 60
 QY 61 CGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 120
 Db 61 CGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 120
 QY 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGACAATGGCTCATCAGCGGAAG 180
 Db 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGACAATGGCTCATCAGCGGAAG 180
 QY 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
 Db 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
 QY 241 CAGAGTGAAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300
 Db 241 CAGAGTGAAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300
 QY 301 ACCCTGCAAGTGCAGCTGCCACACCCAGGAGACCTTCTGCTCTGGTCTCTTAAAGCAC 360
 Db 301 ACCCTGCAAGTGCAGCTGCCACACCCAGGAGACCTTCTGCTCTGGTCTCTTAAAGCAC 360
 QY 361 AGCTCCCTGGGCTGCCAGCGCACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420
 Db 361 AGCTCCCTGGGCTGCCAGCGCACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420
 QY 421 ATCTTTGAACGTGACAGAGACCCAGGAGGAGTACTACTTCCATATTCAGAGCGAACGC 480
 Db 421 ATCTTTGAACGTGACAGAGACCCAGGAGGAGTACTACTTCCATATTCAGAGCGAACGC 480
 QY 481 GCCAACTACACAGTACTTTCACAGTGAATGTAAGAGATACACAGTGTATGCTAAAGG 540
 Db 481 GCCAACTACACAGTACTTTCACAGTGAATGTAAGAGATACACAGTGTATGCTAAAGG 540
 QY 541 AGACCTTACTTTAGGAAGATGAAACACAGGATGCACTGCTCTGCACTCTCCGAGGCTGT 600
 Db 541 AGACCTTACTTTAGGAAGATGAAACACAGGATGCACTGCTCTGCACTCTCCGAGGCTGT 600
 QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 Db 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 QY 661 GGCCTGTGCTGTCAGAGGAGGAGAAAGGTACTTTCATGATGTTTTCGGAACAGACATC 720
 Db 661 GGCCTGTGCTGTCAGAGGAGGAGAAAGGTACTTTCATGATGTTTTCGGAACAGACATC 720
 QY 721 ACATGCTGTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTACCATAGATCTA 780
 Db 721 ACATGCTGTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTACCATAGATCTA 780
 QY 781 AACCAAGGCTCTCAGAGCACACTGCCCCAGTATTCTCTGAAAGTGGGGAACCTTGTGG 840
 Db 781 AACCAAGGCTCTCAGAGCACACTGCCCCAGTATTCTCTGAAAGTGGGGAACCTTGTGG 840
 QY 841 ATCAGTGTGAAGGCGCATCCATGTGAACCATGGATTCGGGCTCACTCGGAGCTGGAAGAC 900
 Db 841 ATCAGTGTGAAGGCGCATCCATGTGAACCATGGATTCGGGCTCACTCGGAGCTGGAAGAC 900
 QY 901 AAAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGACGACCATG 960
 Db 901 AAAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGACGACCATG 960
 QY 961 ATTCGGATTCCTTGGCCCTTGTGCTCTCCGTTGGGAAGGAACGACACCGGATATTACACC 1020
 Db 961 ATTCGGATTCCTTGGCCCTTGTGCTCTCCGTTGGGAAGGAACGACACCGGATATTACACC 1020
 QY 1021 TGCTCTTCTCAAGACACCCAGCCAGTACGCGTTGGTGACCATCTCTAGAAAAAGGTTT 1080
 Db 1021 TGCTCTTCTCAAGACACCCAGCCAGTACGCGTTGGTGACCATCTCTAGAAAAAGGTTT 1080
 QY 1081 ATAACGCTACCAGCTCGCAAGAAGAGTATGAAATGACCCGTCAGAAAAGTTCTGCTTC 1140

1081 ATAAACGCTACCAAGCTCGCAAGAGATGAAATGACCCGTACGAAAAGTTCTGCTTC 1140
1141 TCAGTCAGGTTTAAACGCTACCCAGAAATCCGATGCGAGTGAGTCTTCTCAAGCCTCA 1200
1141 TCAGTCAGGTTTAAACGCTACCCAGAAATCCGATGCGAGTGAGTCTTCTCAAGCCTCA 1200
1201 TTTCTCTGTAACACAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
1201 TTTCTCTGTAACACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
1261 AAGAAACAGCCAGGAGATACATATTTATGAGAAATATGATGCGCCAGTTTCAACAAA 1320
1261 AAGAAACAGCCAGGAGATACATATTTATGAGAAATATGATGCGCCAGTTTCAACAAA 1320
1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCGCTACGCCAGCCAG 1380
1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCGCTACGCCAGCCAG 1380
1381 GCGTCTCTTCTCTGATGGCTACCGCTACCCCTCTTGGACCTGGAAGAAGTTCGGGAC 1440
1381 GCGTCTCTTCTCTGATGGCTACCGCTACCCCTCTTGGACCTGGAAGAAGTTCGGGAC 1440
1441 AAATCTCCAAATGTCACGAGGAAATCCCAAGAGGATTTGGAATATAAAGGCTAACAGA 1500
1441 AAATCTCCAAATGTCACGAGGAAATCCCAAGAGGATTTGGAATATAAAGGCTAACAGA 1500
1501 AAAGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCGGGAAAGG 1560
1501 AAAGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCGGGAAAGG 1560
1561 CTTCTGGTCAATGCTGTGCTACAAATCTATGGCACGCTCTGCGAACAATCTTTTA 1620
1561 CTTCTGGTCAATGCTGTGCTACAAATCTATGGCACGCTCTGCGAACAATCTTTTA 1620
1621 AACTCACCAGGCCCCCTTCCCTTTTCAATCAAGAACACATCTCTTATCGACCACTTGG 1680
1621 AACTCACCAGGCCCCCTTCCCTTTTCAATCAAGAACACATCTCTTATGCGACCACTGG 1680
1681 CTCGTCTCCCTTCATTTGTTCTCATTTGTTGATGTCGCCACAAATACAAAAGCAA 1740
1681 CTCGTCTCCCTTCATTTGTTCTCATTTGTTGATGTCGCCACAAATACAAAAGCAA 1740
1741 TTTAGTACGAGAGTCAGCTGAGATGATCCAGGTGACTGGCCCCCTGTGATACAGGATAC 1800
1741 TTTAGTACGAGAGTCAGCTGAGATGATCCAGGTGACTGGCCCCCTGTGATACAGGATAC 1800
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAC 1860
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAC 1860
1861 TTTAGAGTTGGAGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTC 1920
1861 TTTAGAGTTGGAGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTC 1920
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
1921 TATGCAATAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
2161 TTTACAGGACATGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCCCTACTTTC 2220

2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCCTACTTTC 2220
2221 CAGGCACATTTCAAAATTCAGCATGCTGTTTACGAGAAAGTTTCACTTACACCGCCCTTG 2280
2221 CAGGCACATTTCAAAATTCAGCATGCTGTTTACGAGAAAGTTTCACTTACACCGCCCTTG 2280
2281 GATCAGCTCTCAGGTTCAATGGAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 2340
2281 GATCAGCTCTCAGGTTCAATGGAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 2340
2341 AACCAAGAGGCTGCGCAGAAAGAGAGGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400
2341 AACCAAGAGGCTGCGCAGAAAGAGAGGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400
2401 CTTCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTTCAAGTCTGT 2460
2401 CTTCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTTCAAGTCTGT 2460
2461 GTCCACAGAGACTGCGCAGCAGGAATGTTTGGTCAACCCACCGGAAGTGGTGAAGATC 2520
2461 GTCCACAGAGACTGCGCAGCAGGAATGTTTGGTCAACCCACCGGAAGTGGTGAAGATC 2520
2521 TGTGACTTTGGACTGCGCCGAGACATCTTGGAGCTTCCAGCTACGCTGCTGAGGGCAAC 2580
2521 TGTGACTTTGGACTGCGCCGAGACATCTTGGAGCTTCCAGCTACGCTGCTGAGGGCAAC 2580
2581 GCACGCTCCGCTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAATC 2640
2581 GCACGCTCCGCTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAATC 2640
2641 AAGAGTGAAGTCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGCAAC 2700
2641 AAGAGTGAAGTCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGCAAC 2700
2701 CTTTACCTCTGGACTTCTGTCGACGCTAATCTATAAACTGATTCAGAGTGGATTTAAA 2760
2701 CTTTACCTCTGGACTTCTGTCGACGCTAATCTATAAACTGATTCAGAGTGGATTTAAA 2760
2761 ATGAGCAGCCTTCTATGCCACAGAGGATATCTTGTATGTAATCTGCTGCTGGCT 2820
2761 ATGAGCAGCCTTCTATGCCACAGAGGATATCTTGTATGTAATCTGCTGCTGGCT 2820
2821 TTTGACTCAAGGAAGCGGCTCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2880
2821 TTTGACTCAAGGAAGCGGCTCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2880
2881 GCAGAGCGAAGAAAGCATGTATCAGAACTCCATCCATCTACCAAAACAGCGGCCCT 2940
2881 GCAGAGCGAAGAAAGCATGTATCAGAACTCCATCCATCTACCAAAACAGCGGCCCT 2940
2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCACAGCGCCAGTCCAGAGGTTGAGAGTTTCAAGAA 3000
2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCACAGCGCCAGTCCAGAGGTTGAGAGTTTCAAGAA 3000
3001 AGAAGTTAGCGAGAGCGCTTGGACCCCGCCCTAGCAGGCTGTAGCCGAGAGCA 3060
3001 AGAAGTTAGCGAGAGCGCTTGGACCCCGCCCTAGCAGGCTGTAGCCGAGAGCA 3060
3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGCTTGTCTGCTGAGACTTTTCTCT 3120
3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGGCTTGTCTGCTGAGACTTTTCTCT 3120
3121 AGATGCTGTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180
3121 AGATGCTGTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180
3181 GCGGAGAGCAATAATAGACTTGTGGTGACCCCGCTACCCCTGGGGCTTTTCCAG 3240
3181 GCGGAGAGCAATAATAGACTTGTGGTGACCCCGCTACCCCTGGGGCTTTTCCAG 3240
3241 AGCTTCAGGGAAGCAATGATCTGAAATATAGTATTTCTGTAATACGTGAACAA 3300
3241 AGCTTCAGGGAAGCAATGATCTGAAATATAGTATTTCTGTAATACGTGAACAA 3300

QY 3301 ACCAAACCGTTTTTTTGGCTAAGGAAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA 3360
|||||
Db 3301 ACCAAACCGTTTTTTTGGCTAAGGAAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA 3360
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGTATATATTTATGGATGGAATAAACTTTC 3420
|||||
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGTATATATTTATGGATGGAATAAACTTTC 3420
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
|||||
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 7

AAQ79068 standard; cDNA; 3453 BP.

XX AC AAQ79068;
XX 04-JUL-1995 (first entry)
XX Mouse flk-2 cDNA.
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell; ds.
XX Mus sp.

Location/Qualifiers

58..3039

/*tag= a

58..138

/*tag= b

139..3036

/*tag= c

US5367057-A.

22-NOV-1994.

02-APR-1991; 91US-0679666.

02-APR-1991; 91US-0679666.

28-JUN-1991; 91US-0728913.

15-NOV-1991; 91US-0793065.

24-DEC-1991; 91US-0813593.

26-JUN-1992; 92US-0906397.

12-NOV-1992; 92US-0975049.

19-NOV-1992; 92US-0977451.

30-APR-1993; 93US-0055269.

(UYPR-) UNIV PRINCETON.

Lemischka IR;

WPI; 1995-005894/01.

P-PSDB; AAR67535.

XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
PT proliferation and/or stimulation of primitive mammalian
PT hematopoietic stem cells in vitro or in vivo.

XX Disclosure; Fig. 1A-1F; 69pp; English.

XX cDNAs encoding receptor protein tyrosine-kinases; mouse fetal liver
CC kinase-2 (flk-2); human flk-2 and mouse flk-1 are given in AAQ79068-70,
CC respectively, and the deduced amino acid sequences in AAR67535-37,
CC respectively.

XX Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;

XX Query Match 100.0%; Score 3451.4; DB 16; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCGCTGCTACCGCGCGCTCCGGAGGCGCATCGCGCGCTTGGCGAGCGCAGCGACCGG 60
Db 1 GGGCGCTGCTACCGCGCGCTCCGGAGGCGCATCGCGCGCTTGGCGAGCGCAGCGACCGG 60
QY 61 CGGCTGCTGCTGCTTGTGTTTGTGTCAGTAAATGATTTCTTGAGACCGCTTACAAACCAAG 120
Db 61 CGGCTGCTGCTGCTTGTGTTTGTGTCAGTAAATGATTTCTTGAGACCGCTTACAAACCAAG 120
QY 121 CTGCTGTGATCAAGTGTGTTTTTAATAGTCATGAGAAATGCTCATCAGCGGGAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTTTAATAGTCATGAGAAATGCTCATCAGCGGGAAG 180
QY 181 CCATCATCGTACCGAATGTCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
Db 181 CCATCATCGTACCGAATGTCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
QY 241 CAGAGTGAAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300
Db 241 CAGAGTGAAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300
QY 301 ACCCTGCAAGTCAGCTGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTCAGCTGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGTCGCCAGCGCCTTTGATTTACAAAACAGAGAAATCTTCCATGCC 420
Db 361 AGCTCCCTGGGTCGCCAGCGCCTTTGATTTACAAAACAGAGAAATCTTCCATGCC 420
QY 421 ATCTTGAAGTCAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACGC 480
Db 421 ATCTTGAAGTCAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACGC 480
QY 481 GCCAATACACAGTACTGTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAATACACAGTACTGTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540
QY 541 AGACCTTACTTTAGGAGATGGAACACAGGATGCACCTGCTGCTCATCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAGATGGAACACAGGATGCACCTGCTGCTCATCTCCGAGGGTGT 600
QY 601 CCGGAGCCCTGCTGGAGTGGGTGCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA 660
Db 601 CCGGAGCCCTGCTGGAGTGGGTGCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA 660
QY 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCAATGAGTTGTTCCGAAACAGACATC 720
Db 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCAATGAGTTGTTCCGAAACAGACATC 720
QY 721 AGATGCTGTGCTAGAAATGCACCTGGCGCGAATGCACCAAGCTGTTCCACCATAGATCTA 780
Db 721 AGATGCTGTGCTAGAAATGCACCTGGCGCGAATGCACCAAGCTGTTCCACCATAGATCTA 780
QY 781 AACACGCTCCTCAGAGCACACTGCCCGAGTTTCTGAAAGTGGGGGAAACCTTGTGG 840
Db 781 AACACGCTCCTCAGAGCACACTGCCCGAGTTTCTGAAAGTGGGGGAAACCTTGTGG 840
QY 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACTCCAGAGTGGAGAC 900
Db 841 ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACTCCAGAGTGGAGAC 900
QY 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGAGACCATG 960
Db 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGAGACCATG 960
QY 961 ATTCGATTCCTCTTGGCTTTGTCTCCGTCGGAAGGAGACGACCGGATATTACACC 1020
Db 961 ATTCGATTCCTCTTGGCTTTGTCTCCGTCGGAAGGAGACGACCGGATATTACACC 1020
QY 1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGTTGACCTCTAGAAAAAGGTTT 1080
|||||

QY 3241 AGCTTGGGGAAAGCCATGATCTGAATATAGTATATATCTTCTTAAATAGCTGAACAA 3300
 Db |||||||
 QY 3301 ACCAAACCCGTTTTCGTAAGGAAAGCTAAATATGATTTTAAATATCTATGTTTAA 3360
 Db |||||||
 QY 3361 AATACTATGTAACTTTTTCATCTATTAGTGATATATTTATGGATGGAATAAACTTTC 3420
 Db |||||||
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAA 3453
 Db |||||||
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAA 3453

RESULT 8

AAQ29954
 ID AAQ29954 standard; cDNA; 3453 BP.

XX
 AC AAQ29954;

XX 15-MAR-1993 (first entry)

XX Murine flk-2 cDNA sequence.

XX Thymidine kinase; TK; haematopoietic; stem cells; proliferation;
 KW differentiation; progenitor cells; foetal liver kinase; ss.

XX Mus musculus.

PH Key Location/Qualifiers
 FT CDS 31..3009
 FT /*tag= a

XX WO9217486-A.

XX 15-OCT-1992.

XX 02-APR-1992; 92WO-US02750.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1992-366185/44.

XX P-PSDB; AAR28038.

XX Stimulating proliferation and/or differentiation of primitive
 PT mammalian haematopoietic stem cells - using ligand that binds
 PT thymidine kinase and flk-1 and flk-2

XX Claim 6; Fig 1a; 94pp; English.

XX The murine foetal liver kinase (flk)-2 clone was isolated by
 CC standard PCR techniques from stem-cell receptor-contg. tissue cDNA
 CC libraries. Suitable tissues include foetal liver, spleen or thymus
 CC cells or adult marrow or brain cells. The PCR primers used are based
 CC on known sections of the flk-2 gene. The murine flk-2 clone may be
 CC used in a vector to transform haematopoietic cells. The thymidine
 CC kinase encoded by flk-2 is expressed in primitive but not mature
 CC haematopoietic cells. Ligand binding to the TK may be prepd. which
 CC can stimulate proliferation and/or differentiation of primitive
 CC haematopoietic cells in vivo. The ligands can stimulate the
 CC proliferation of additional primitive stem cells, differentiation into
 CC more mature progenitor cells, or both.
 CC See also AAQ29955-7.

XX Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;
 SQ Query Match 99.9%; Score 3448.2; DB 13; Length 3453;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGCGTTGGCGAGCGACGCGCGG 60
 Db |||||||
 QY 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGCGTTGGCGAGCGACGCGCGG 60
 Db |||||||
 QY 61 CGGCTGCTGCTGCTTGTGTCAGTAATGATTCTTGAGACCGCTTACAACCAAGAC 120
 Db |||||||
 QY 61 CGGCTGCTGCTGCTTGTGTCAGTAATGATTCTTGAGACCGCTTACAACCAAGAC 120
 Db |||||||
 QY 121 CTGCGCTGTCATCAAGTGTCTTTTAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180
 Db |||||||
 QY 121 CTGCGCTGTCATCAAGTGTCTTTTAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180
 Db |||||||
 QY 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACTCCAGTGTACCCCGAGGCGC 240
 Db |||||||
 QY 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACTCCAGTGTACCCCGAGGCGC 240
 Db |||||||
 QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGCCACCGTGGAGGTGCCCGAGTCTGGTCCATC 300
 Db |||||||
 QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGCCACCGTGGAGGTGCCCGAGTCTGGTCCATC 300
 Db |||||||
 QY 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGAGCTTTCCTGCCCTCTGGGTCTTTAAGCAC 360
 Db |||||||
 QY 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGAGCTTTCCTGCCCTCTGGGTCTTTAAGCAC 360
 Db |||||||
 QY 361 AGCTCCCTGGCTGCCAGCGGCTTTGATTTACAAAACAGAGAGAAATCGTTCCATGGCC 420
 Db |||||||
 QY 361 AGCTCCCTGGCTGCCAGCGGCTTTGATTTACAAAACAGAGAGAAATCGTTCCATGGCC 420
 Db |||||||
 QY 421 ATCTTTAAGCTGACAGAGACCCAGGAGGAGAAATACCTACTTCATATTCAGCGGAACGC 480
 Db |||||||
 QY 421 ATCTTTAAGCTGACAGAGACCCAGGAGGAGAAATACCTACTTCATATTCAGCGGAACGC 480
 Db |||||||
 QY 481 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540
 Db |||||||
 QY 481 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540
 Db |||||||
 QY 541 AGACCTTACTTTAGGAAGATGAAAACCCAGGATGCTGCTGCTCATCTCCGAGGGTGT 600
 Db |||||||
 QY 541 AGACCTTACTTTAGGAAGATGAAAACCCAGGATGCTGCTGCTCATCTCCGAGGGTGT 600
 Db |||||||
 QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTGTGAGTCCCAAGGAAAGCTGTAAGAAGAA 660
 Db |||||||
 QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTGTGAGTCCCAAGGAAAGCTGTAAGAAGAA 660
 Db |||||||
 QY 661 GGGCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCATGAGTGTTCGGAACAGACATC 720
 Db |||||||
 QY 661 GGGCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCATGAGTGTTCGGAACAGACATC 720
 Db |||||||
 QY 721 AGATGCTGTGCTAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCAACCATAGATCTA 780
 Db |||||||
 QY 721 AGATGCTGTGCTAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCAACCATAGATCTA 780
 Db |||||||
 QY 781 AACCAAGGCTCCTCAGAGCACACTGCCCCAGTATTCTTGAAAGTGGGGGAACCCCTTGG 840
 Db |||||||
 QY 781 AACCAAGGCTCCTCAGAGCACACTGCCCCAGTATTCTTGAAAGTGGGGGAACCCCTTGG 840
 Db |||||||
 QY 841 ATCAGTGTGAAGCCCATCATGTGAACCATGGATTTCGGGCTCACCTGGGAGCTGGAGAC 900
 Db |||||||
 QY 841 ATCAGTGTGAAGCCCATCATGTGAACCATGGATTTCGGGCTCACCTGGGAGCTGGAGAC 900
 Db |||||||
 QY 901 AAGACCCCTGGAGGAGGCGAGCTACTTTGAGATCAGTACCTACTCCACAAACAGGACCATG 960
 Db |||||||
 QY 901 AAGACCCCTGGAGGAGGCGAGCTACTTTGAGATCAGTACCTACTCCACAAACAGGACCATG 960
 Db |||||||
 QY 961 ATTCCGATCTCTTGGCTTTGTGCTTCCCTGGGGAAGGAACGACACCCGGATATTACACC 1020
 Db |||||||

Db 961 AFTCGGATTCCTTGGCCCTTTGTGCTTCCGTGGGAAGGAACGACACCCGGATATTACACC 1020
Qy 1021 TCCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGTTT 1080
Db 1021 TGCTCTTCTTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGTTT 1080
Qy 1081 ATAAACGCTACAGCTCGCAAGAAGATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
Db 1081 ATAAACGCTACAGCTCGCAAGAAGATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAACGCTACCCAGATCCGATCCGATCGATGATCTTCTCAAGCCTCA 1200
Db 1141 TCAGTCAGGTTTAAACGCTACCCAGATCCGATCCGATCGATGATCTTCTCAAGCCTCA 1200
Qy 1201 TTTCCCTGTGAACAGAGGCTCGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260
Db 1201 TTTCCCTGTGAACAGAGGCTCGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260
Qy 1261 AAGAACAAAGCCAGGAGGTACATATTCTATGCAGAAATGATGAGCCCGAGTTTCAACAAA 1320
Db 1261 AAGAACAAAGCCAGGAGGTACATATTCTATGCAGAAATGATGAGCCCGAGTTTCAACAAA 1320
Qy 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTACGCCAGCCAG 1380
Db 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTACGCCAGCCAG 1380
Qy 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAAGTGTTCGGAC 1440
Db 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAAGTGTTCGGAC 1440
Qy 1441 AAATCTCCCAATTGACAGAGGAAATCCAGAAAGGATTTGGAATAAAAAAGGCTAACAGA 1500
Db 1441 AAATCTCCCAATTGACAGAGGAAATCCAGAAAGGATTTGGAATAAAAAAGGCTAACAGA 1500
Qy 1501 AAAGTGTGTTGGCCAGTGGGTGCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560
Db 1501 AAAGTGTGTTGGCCAGTGGGTGCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560
Qy 1561 CTCTCGTCAAAATGCTGCGTACAAATCTATGCGCAGCTCTTGGGAACCATCTTTTFA 1620
Db 1561 CTCTCGTCAAAATGCTGCGTACAAATCTATGCGCAGCTCTTGGGAACCATCTTTTFA 1620
Qy 1621 AACTCACCAGGCCCTTCCCTTTCATCCAAAGAACATCTCTTATGCGACCATTTGG 1680
Db 1621 AACTCACCAGGCCCTTCCCTTTCATCCAAAGAACATCTCTTATGCGACCATTTGG 1680
Qy 1681 CTCTCTCCCTTCATTTGTTCTCATTTGTTGATGTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTCTCCCTTCATTTGTTCTCATTTGTTGATGTGCCACAAATACAAAAGCAA 1740
Qy 1741 TTTAGGTACGAGTCAAGTCAGATGATCCAGGTGACTGGGCCCTGGATACAGGTAC 1800
Db 1741 TTTAGGTACGAGTCAAGTCAGATGATCCAGGTGACTGGGCCCTGGATACAGGTAC 1800
Qy 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGTGAAGAC 1860
Qy 1861 TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGTTGATGAACCCAGGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGTTGATGAACCCAGGCC 1920
Qy 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGGGTGAAGATGTCTAAAGAGAAA 1980
Db 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGGGTGAAGATGTCTAAAGAGAAA 1980
Qy 1981 GCTGACAGTGTGAAGAAAGGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGTGTGAAGAAAGGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100

Qy 2101 ATTTTGAATATTGTTGCTATGCTGACCTCCTCAACTACCTAAGAACTAAAAAGAGAAG 2160
Db 2101 ATTTTGAATATTGTTGCTATGCTGACCTCCTCAACTACCTAAGAACTAAAAAGAGAAG 2160
Qy 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTCAAGTCTTTACCCTACTTTC 2220
Db 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTCAAGTCTTTACCCTACTTTC 2220
Qy 2221 CAGGACATTTCAAAATTCAGCATGCCTGGTTTCAGAGAAAGTTTCAGTTTACACCCGCCCTTG 2280
Db 2221 CAGGACATTTCAAAATTCAGCATGCCTGGTTTCAGAGAAAGTTTCAGTTTACACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCTCAAGATGAGATTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCTCAAGATGAGATTGAATATGAA 2340
Qy 2341 AACCAAGAGAGGCTGCGCAGAAAGAGGAGAAATTTGAACGTGCTGACGTTTGAAGAC 2400
Db 2341 AACCAAGAGAGGCTGCGCAGAAAGAGGAGAAATTTGAACGTGCTGACGTTTGAAGAC 2400
Qy 2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGT 2460
Db 2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGT 2460
Qy 2461 GTCCACAGAGCTGCGCAGCCAGGAATGTGTTGTCACCCACCGGAAGGTGGTGAAGATC 2520
Db 2461 GTCCACAGAGCTGCGCAGCCAGGAATGTGTTGTCACCCACCGGAAGGTGGTGAAGATC 2520
Qy 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580
Qy 2581 GCAGGCTCCGCTGAAGTGGATGGCACCCGAGAGCTTATTGAAGGGATCTACACAATC 2640
Db 2581 GCAGGCTCCGCTGAAGTGGATGGCACCCGAGAGCTTATTGAAGGGATCTACACAATC 2640
Qy 2641 AAGAGTGAAGTCTGCTGCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Db 2641 AAGAGTGAAGTCTGCTGCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Qy 2701 CCTTACCTCGGCAATTCCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTCGGCAATTCCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760
Qy 2761 ATGGACAGCCATTCATGCCACAGAGGATATCTTGTATGCAATCCTGCTGGGCT 2820
Db 2761 ATGGACAGCCATTCATGCCACAGAGGATATCTTGTATGCAATCCTGCTGGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCATCTTTCCTCCCAACCTGACTTTCTTTTAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTTTCCTCCCAACCTGACTTTCTTTTAGGATGTCAGCTG 2880
Qy 2881 GCAGAGCGCAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGCGCAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940
Qy 2941 CAGCAGAGAGGGGGTCTAGAGCCCATGTCGCCACAGCCAGCGGATTCACAGAGAA 3000
Db 2941 CAGCAGAGAGGGGGTCTAGAGCCCATGTCGCCACAGCCAGCGGATTCACAGAGAA 3000
Qy 3001 AGAAGTTACGAGGAGGCCCTTGGACCCGCCCTAGCAGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTACGAGGAGGCCCTTGGACCCGCCCTAGCAGCTGTAGACCGCAGAGCCA 3060
Qy 3061 AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGCTTGTCTTCCGCTGGACTTTTCTCT 3120
Db 3061 AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTTGTCTTCCGCTGGACTTTTCTCT 3120
Qy 3121 AGATTGCTGTGCTTACTTCAAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATTGCTGTGCTTACTTCAAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180

QY 3181 GCGGAGAGCAATAATGAGACTTCTGGTGAGCCCGCTACCTGGGGGCTTTCCACG 3240
 DB 3181 GCGGAGAGCAATAATGAGACTTCTGGTGAGCCCGCTACCTGGGGGCTTTCCACG 3240
 QY 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
 DB 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
 QY 3301 ACCAACCCTGTTTCTGCTAAGGGAACCTAAATATGATTTTAAAAATCTATGTTTAA 3360
 DB 3301 ACCAACCCTGTTTCTGCTAAGGGAACCTAAATATGATTTTAAAAATCTATGTTTAA 3360
 QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGAAATAAATTTTC 3420
 DB 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGAAATAAATTTTC 3420
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
 DB 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 9

AAQ40914
 ID AAQ40914 standard; cDNA; 3453 BP.

XX
 AC AAQ40914;

XX 19-OCT-1993 (first entry)

XX Murine flk-2 cDNA.

XX Murine: receptor; protein; tyrosine kinase; ptk; flk-2; primitive;
 KW hematopoietic cell; mature; family; conserved; region;
 KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;
 KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;
 KW hierarchy; transduction; T-lymphoid; lineage; ss.

XX Mus musculus.

XX Key Location/Qualifiers
 FH CDS 31..3009

FT /*tag= a

FT sig_peptide 31..111

FT /*tag= b

FT /note= "Hydrophobic leader sequence"

FT mat_peptide 112..3006

FT /*tag= c

XX WO9310136-A.

XX 27-MAY-1993.

XX 16-NOV-1992; 92WO-US09893.

XX 15-NOV-1991; 91US-0793065.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-182479/22.

XX P-PSDB; AAR37502.

XX Totipotent haematopoietic stem cell receptors, their ligands and
 PT DNA sequences - for treating anemia(s) and bone marrow damage
 PT due to e.g. cancer chemotherapy or radiotherapy

XX Claim 6; Fig 1a; 127pp; English.

XX This sequence encodes the murine receptor protein tyrosine kinase
 CC (PTK), flk-2. This nucleic acid is expressed in primitive hemo-
 CC poietic cells and not in mature hematopoietic cells. Members of
 CC this family of ptk's can be recognised by the conserved amino acid

CC regions in the catalytic domain. This family of ptk's also contains
 CC c-kit. These new receptors are termed fetal liver kinases (flk's)
 CC after the tissue in which they were discovered. flk-2 is also
 CC expressed in fetal spleen, fetal thymus, adult brain and adult
 CC bone marrow. flk-2 is expressed in individual multipotential CFU-
 CC Blast colonies capable of generating numerous multilineage colonies
 CC upon replating. It is likely therefore, that flk-2 is expressed in
 CC the entire primitive portion of the hematopoietic hierarchy. This is
 CC consistent with flk-2 being important in transducing putative self-
 CC renewal signals from the environment. flk-2 is the first receptor
 CC ptk known to be expressed in the T-lymphoid lineage.

XX Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;

Query Match 99.8%; Score 3446.6; DB 14; Length 3453;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGGCTGCTGCTACCGCGGCTCCGGAGGCCATCGGGCGTTGGCGCAGCGCAGCGACCGG 60

DB 1 GCGGCTGCTGCTACCGCGGCTCCGGAGGCCATCGGGCGTTGGCGCAGCGCAGCGACCGG 60

QY 61 CGGCTGCTGCTGCTGTTGTTTGTCAATGATCTTGTGAGACCCGTTACAAACCAAGAC 120

DB 61 CGGCTGCTGCTGCTGTTGTTTGTCAATGATCTTGTGAGACCCGTTACAAACCAAGAC 120

QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATAGCGGAAAG 180

DB 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATAGCGGAAAG 180

QY 181 CCATCATCTACCGAATGTGCGAGATCCCAAGAGACTCCAGTGTACCCCGAGCGC 240

DB 181 CCATCATCTACCGAATGTGCGAGATCCCAAGAGACTCCAGTGTACCCCGAGCGC 240

QY 241 CAGAGTGAAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGAGTCTGGTCCATC 300

DB 241 CAGAGTGAAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGAGTCTGGTCCATC 300

QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGAGACTTCTCTGCTCTGGGTCTTTAAGCAC 360

DB 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGAGACTTCTCTGCTCTGGGTCTTTAAGCAC 360

QY 361 AGCTCCCTGGCTGCGCAGCGCACCTTTGATTTACAAACAGAGAAATCGTTTCATGCC 420

DB 361 AGCTCCCTGGCTGCGCAGCGCACCTTTGATTTACAAACAGAGAAATCGTTTCATGCC 420

QY 421 ATCTTGAAGTGCACAGACCCAGCGAGAGATACCTACTCCATATTCAGAGCGAAGC 480

DB 421 ATCTTGAAGTGCACAGACCCAGCGAGAGATACCTACTCCATATTCAGAGCGAAGC 480

QY 481 GCCAATACACAGTACTGTTACAGTGAATGTAAAGAGATACAGCTGTATGCTAAGG 540

DB 481 GCCAATACACAGTACTGTTACAGTGAATGTAAAGAGATACAGCTGTATGCTAAGG 540

QY 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCATGCTCTGCTCGAGGGTGT 600

DB 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCATGCTCTGCTCGAGGGTGT 600

QY 601 CCGAGCCCACTGTGGAGTGGTCTGCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660

DB 601 CCGAGCCCACTGTGGAGTGGTCTGCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660

QY 661 GGCCCTGCTGCTTCAGAAAGAGGAGAAAGTACTTTCATGAGTGTTCGGAACAGATC 720

DB 661 GGCCCTGCTGCTTCAGAAAGAGGAGAAAGTACTTTCATGAGTGTTCGGAACAGATC 720

QY 721 AGATGCTGTGCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTACCATAGATCTA 780

DB 721 AGATGCTGTGCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTACCATAGATCTA 780

QY 781 AACCAAGCTCCTCAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGAAACCTTGTGG 840

DB 781 AACCAAGCTCCTCAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGAAACCTTGTGG 840

QY 841 ATCAGGTCTAAGGCCATCCATGCTGAACCATGGATTTCGGGCTCACCTGGAGCTGGAAGAC 900
DB 841 ATCAGGTCTAAGGCCATCCATGCTGAACCATGGATTTCGGGCTCACCTGGAGCTGGAAGAC 900
QY 901 AAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAAACAGACCAATG 960
DB 901 AAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAAACAGACCAATG 960
QY 961 ATTCCGATTCCTTTGGCCCTTTGCTCTCCGTGGGAAGGAACGACACCGGATATTACACC 1020
DB 961 ATTCCGATTCCTTTGGCCCTTTGCTCTCCGTGGGAAGGAACGACACCGGATATTACACC 1020
QY 1021 TGCTCTCTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCATCTAGAAAAAGGGTTT 1080
DB 1021 TGCTCTCTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCATCTAGAAAAAGGGTTT 1080
QY 1081 ATAAACGCTACAGCTCGCAAGAAGATGATGAATTTGACCCGTCAGAAAAGTTCTGCTTC 1140
DB 1081 ATAAACGCTACAGCTCGCAAGAAGATGATGAATTTGACCCGTCAGAAAAGTTCTGCTTC 1140
QY 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCGATGCGATCTCTCAAGCCCTCA 1200
DB 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCGATGCGATCTCTCTCAAGCCCTCA 1200
QY 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
DB 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY 1261 AAGAACAGCCAGGAGAGTACATATTCATGCAAAAAATGATGACGCCAGTTTCAACAAA 1320
DB 1261 AAGAACAGCCAGGAGAGTACATATTCATGCAAAAAATGATGACGCCAGTTTCAACAAA 1320
QY 1321 ATGTTCAAGCTGAATATAGAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
DB 1321 ATGTTCAAGCTGAATATAGAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
QY 1381 GCGTCTCTCTCTGATGGCTACCCGCTACCCCTTGGAGCTGGAAGAGTTTGGGAC 1440
DB 1381 GCGTCTCTCTCTGATGGCTACCCGCTACCCCTTGGAGCTGGAAGAGTTTGGGAC 1440
QY 1441 AAATCTCCCAATTCACGAGGAGAAATCCAGAGGAGTTTGAATATAAAGGCTAACAGA 1500
DB 1441 AAATCTCCCAATTCACGAGGAGAAATCCAGAGGAGTTTGAATATAAAGGCTAACAGA 1500
QY 1501 AAAGTGTGTGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCGGGAAGGG 1560
DB 1501 AAAGTGTGTGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCGGGAAGGG 1560
QY 1561 CTTCTGGTCAATGCTGTGGGTACAAATTCATGGGCAGCTCTGGCAACCATCTTTTA 1620
DB 1561 CTTCTGGTCAATGCTGTGGGTACAAATTCATGGGCAGCTCTGGCAACCATCTTTTA 1620
QY 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTTATGCGACCATTTGG 1680
DB 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTTATGCGACCATTTGG 1680
QY 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
DB 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
QY 1741 TTTAGGTACGAGATCAGTCAGATGATCAGGTGACTGGGCCCTGGATTAACAGTAC 1800
DB 1741 TTTAGGTACGAGATCAGTCAGATGATCAGGTGACTGGGCCCTGGATTAACAGTAC 1800
QY 1801 TTCTACGTTGACTTCAGGAGTATCAATATGACCTTAAGTGGAGTTCCTCCGAGAGAAC 1860
DB 1801 TTCTACGTTGACTTCAGGAGTATCAATATGACCTTAAGTGGAGTTCCTCCGAGAGAAC 1860
QY 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGTTGATTAACGCCAGGCC 1920
DB 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGTTGATTAACGCCAGGCC 1920

QY 1921 TATGCATTAGTAAACGGGAGCTCTCAATTCAGGTGGGTCAGATGCTAAAGAGAAA 1980
DB 1921 TATGCATTAGTAAACGGGAGCTCTCAATTCAGGTGGGTCAGATGCTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
DB 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACATGACACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2100
DB 2041 CACATGACACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2100
QY 2101 ATTTTGAATATGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
DB 2101 ATTTTGAATATGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCTACTTTC 2220
DB 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCTACTTTC 2220
QY 2221 CAGGCACATTCNAATTCAGCATGCTGGTTCACGAGAAGTTTCAGTTACACCCGCCCTTG 2280
DB 2221 CAGGCACATTCNAATTCAGCATGCTGGTTCACGAGAAGTTTCAGTTACACCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTTCAATTCCTGAAGATGAGATTGAATGAA 2340
DB 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTTCAATTCCTGAAGATGAGATTGAATGAA 2340
QY 2341 AACCAGAAGAGGCTGGCAGAAAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400
DB 2341 AACCAGAAGAGGCTGGCAGAAAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTCTTGAGTTCAAGTCTGT 2460
DB 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTCTTGAGTTCAAGTCTGT 2460
QY 2461 GTCCACAGAGCTCGCAGCCAGGAGTGTGGTTCACCCAGGGAAGTGGTGAAGATC 2520
DB 2461 GTCCACAGAGCTCGCAGCCAGGAGTGTGGTTCACCCAGGGAAGTGGTGAAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGGAGACATCTCTGAGCGACTCCAGCTACGCTCAGGGGCAAC 2580
DB 2521 TGTGACTTTGGACTGGCCCGGAGACATCTCTGAGCGACTCCAGCTACGCTCAGGGGCAAC 2580
QY 2581 GCACGGCTGCCGCTGAAGTGGATGACCCGAGAGCTTATTTGAAGGATCTACAAATC 2640
DB 2581 GCACGGCTGCCGCTGAAGTGGATGACCCGAGAGCTTATTTGAAGGATCTACAAATC 2640
QY 2641 AAGAGTACGCTGTGGTCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
DB 2641 AAGAGTACGCTGTGGTCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
QY 2701 CCTTACCCTGGCATCTCTGTCAGCGCTAACTTCTATAAACTGATTTCAGAGTGGATTAAA 2760
DB 2701 CCTTACCCTGGCATCTCTGTCAGCGCTAACTTCTATAAACTGATTTCAGAGTGGATTAAA 2760
QY 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATACTTTGTAATGCAATCTCTCTGGGCT 2820
DB 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATACTTTGTAATGCAATCTCTCTGGGCT 2820
QY 2821 TTTGACTCAAGAGGCGGCATCTTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
DB 2821 TTTGACTCAAGAGGCGGCATCTTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
QY 2881 GCAGAGGAGAGAGAGATGATCAGAACATCCATCTTACCAAAACAGGGGCGCCCT 2940
DB 2881 GCAGAGGAGAGAGAGATGATCAGAACATCCATCTTACCAAAACAGGGGCGCCCT 2940
QY 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACACAGCGCCAGGTGAAGATTACAGAGAA 3000
DB 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACACAGCGCCAGGTGAAGATTACAGAGAA 3000
QY 3001 AGAAGTTACGAGGAGGCGCTTGGACCCCGCCCTTAGCAGGCTGTAGACCCGAGAGCCA 3060

QY 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGTAGTACCTACTCTCAACAAACAGGACCATG 960
DB 952 AAAGCCCTGGAGGAGGAGCTACTTTGAGTAGTACCTACTCTCAACAAACAGGACCATG 1011
QY 961 ATTCCGATTTCTTTGGCCCTTTGCTCTCCCGTGGGAAGAACAGACACCGGATATTACACC 1020
DB 1012 ATTCCGATTTCTTTGGCCCTTTGCTCTCCCGTGGGAAGAACAGACACCGGATATTACACC 1071
QY 1021 TCTCTCTCTCAAAACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080
DB 1072 TCTCTCTCTCAAAACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1131
QY 1081 ATAAACGGTACCAGCTCGCAAGAAAGATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
DB 1132 ATAAACGGTACCAGCTCGCAAGAAAGATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1191
QY 1141 TCAGTCAGGTTTAAACGGTACCAGCAAGATCCGATGACGTGATCTTCTCAAGCCTCA 1200
DB 1192 TCAGTCAGGTTTAAACGGTACCAGCAAGATCCGATGACGTGATCTTCTCAAGCCTCA 1251
QY 1201 TTTCTCTTGTAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTCCGATCAT 1260
DB 1252 TTTCTCTTGTAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTCCGATCAT 1311
QY 1261 AAGAAACAAGCCAGGAGTACATATTTCTATGAGAAAATGATGACCCCGAGTTTCAACAAA 1320
DB 1312 AAGAAACAAGCCAGGAGTACATATTTCTATGAGAAAATGATGACCCCGAGTTTCAACAAA 1371
QY 1321 ATGTTACGCTGAATATAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
DB 1372 ATGTTACGCTGAATATAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1431
QY 1381 GGTCTCTGTTCTGATGGCTACCGCTACCTCTTGACCTGGAGAGTGTTCGGAC 1440
DB 1432 GGTCTCTGTTCTGATGGCTACCGCTACCTCTTGACCTGGAGAGTGTTCGGAC 1491
QY 1441 AAATCTCCCAATTTGACGAGGAGAAATCCAGAGAGGTTTGGAAATAAAAGGCTACACGA 1500
DB 1492 AAATCTCCCAATTTGACGAGGAGAAATCCAGAGAGGTTTGGAAATAAAAGGCTACACGA 1551
QY 1501 AAAGTGTTTGGCCAGTGGTGTGAGAGTACTCTAAATATGAGTGAGGCGCGGGAAGGG 1560
DB 1552 AAAGTGTTTGGCCAGTGGTGTGAGAGTACTCTAAATATGAGTGAGGCGCGGGAAGGG 1611
QY 1561 CTTCTGGTCAATGCTGTCGCTACAAATCTATGGGCACGCTTGCGAAACCATCTTTTAA 1620
DB 1612 CTTCTGGTCAATGCTGTCGCTACAAATCTATGGGCACGCTTGCGAAACCATCTTTTAA 1671
QY 1621 AACTCACCAGGCCCTTCCCTTTTATCCCAAGACAACTCTCTTATGCGACCATTTGGG 1680
DB 1672 AACTCACCAGGCCCTTCCCTTTTATCCCAAGACAACTCTCTTATGCGACCATTTGGG 1731
QY 1681 CTTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGGCACAATAACAAAAAGCAA 1740
DB 1732 CTTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGGCACAATAACAAAAAGCAA 1791
QY 1741 TTTAGTAGCAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCTGGATACGAGTAC 1800
DB 1792 TTTAGTAGCAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCTGGATACGAGTAC 1851
QY 1801 TTTCTAGCTTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
DB 1852 TTTCTAGCTTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1911
QY 1861 TTAGAGTTTGGGAAGTCTTGGGCTGTCGCTTTTCGGAGGGTGTGAAACGCCAGGCC 1920
DB 1912 TTAGAGTTTGGGAAGTCTTGGGCTGTCGCTTTTCGGAGGGTGTGAAACGCCAGGCC 1971
QY 1921 TATGGCATTTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAAGAGAAA 1980
DB 1972 TATGGCATTTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAAGAGAAA 2031
QY 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGATGACCCACCTGGGA 2040

DB 2032 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2091
QY 2041 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2100
DB 2092 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2151
QY 2101 ATTTTTGAATATTGTTGCTATGCTGACCTCTCAACTACCTAAAGAACTAAAGAGAGAAG 2160
DB 2152 ATTTTTGAATATTGTTGCTATGCTGACCTCTCAACTACCTAAAGAACTAAAGAGAGAAG 2211
QY 2161 TTTTACAGGACATGACAGAGATTTTTTAAGGAACATAATTTTCACTTCTTACCCTACTTTC 2220
DB 2212 TTTTACAGGACATGACAGAGATTTTTTAAGGAACATAATTTTCACTTCTTACCCTACTTTC 2271
QY 2221 CAGGCACATTCAAATTTCCAGCATGCTGTTTACAGAGAACTTCACTTACACCCGCCCTTG 2280
DB 2272 CAGGCACATTCAAATTTCCAGCATGCTGTTTACAGAGAACTTCACTTACACCCGCCCTTG 2331
QY 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTTCAATTTCAAGATGAGATTGAATATGAA 2340
DB 2332 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTTCAATTTCAAGATGAGATTGAATATGAA 2391
QY 2341 AACCAAGAGAGGCTGCGAGAGAGAGAGAGATTTTGAACGTGTGACGTTTGAAGAC 2400
DB 2392 AACCAAGAGAGGCTGCGAGAGAGAGAGAGATTTTGAACGTGTGACGTTTGAAGAC 2451
QY 2401 CTCCTTTGCTTTCCGTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAAGTCGTGT 2460
DB 2452 CTCCTTTGCTTTCCGTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAAGTCGTGT 2511
QY 2461 GTCACAGAGACCTGGCAGCAGCAATGCTTGGTCAACCCAGGAGGTGGTGAAGATC 2520
DB 2512 GTCACAGAGACCTGGCAGCAGCAATGCTTGGTCAACCCAGGAGGTGGTGAAGATC 2571
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTTCAAGTACGCTGCTGAGGGCAAC 2580
DB 2572 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTTCAAGTACGCTGCTGAGGGCAAC 2631
QY 2581 GCACGGCTCCCGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACATC 2640
DB 2632 GCACGGCTCCCGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACATC 2691
QY 2641 AAGAGTGACCTCTGCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
DB 2692 AAGAGTGACCTCTGCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2751
QY 2701 CTTTACCCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2760
DB 2752 CTTTACCCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2811
QY 2761 ATGAGGAGCCATTTCTATGCCACAGAGGATATACCTTTGATGCAATCTCTGCTGGCT 2820
DB 2812 ATGAGGAGCCATTTCTATGCCACAGAGGATATACCTTTGATGCAATCTCTGCTGGCT 2871
QY 2821 TTTGACTCAAGGAAGGGGCCATCTTCCCAACTGACTTCAATTTTATAGGATGTCAGCTG 2880
DB 2872 TTTGACTCAAGGAAGGGGCCATCTTCCCAACTGACTTCAATTTTATAGGATGTCAGCTG 2931
QY 2881 GCAGAGCGAAGGAAGC-----ATGTPATCAGAACATCCATCC 2917
DB 2932 GCAGAGCGAAGGAAGCGATGTATCAACATGGGTGGCAACGCTCCAGAACATCCATCC 2991
QY 2918 ATCTACAAACAGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCC-CAGTCCGCCACAG 2976
DB 2992 ATCTACAAACAGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCCGCCATCCGCCACAG 3051
QY 2977 CGCCAGGTGAAGATTACACAGAGAAAGATTAGCAGAGAGCGCTTGACCCGCCACCT 3036
DB 3052 GCCCAGGTGAAGATTACACAGAGAAAGATTAGCAGAGAGCGCTTGACCCGCCACCT 3111
QY 3037 AGCAGGCTGTGAGACCGCAGAGCCAAAGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCG 3096

Qy	788	CTCCTCAGACGACACTGCCCCAGTTATTCTGAAAGTGGGGAAACCTCTTGATCAGGT	847
Db	812	CTCCTCAGACGACACTGCCCAATTATTCTTAAAGTAGGGGAACCTTATGATAGAGT	871
Qy	848	GTAAGGCCATCCATGTGAACCATGGATTTCGGGTCACTTCGGAGCTGGAAGACAAAGCCC	907
Db	872	GCAAGCTGTTCATGTGAACCATGGATTTCGGGTCACTTCGGAAATTAGAAAACAAGCAC	931
Qy	908	TGGAGGAGCGAGCTACTTTGAGATGAGTACCTACTCCACAAAACAGACCATGATCGGA	967
Db	932	TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA	991
Qy	968	TTCTCTTTGGCCTTTGTCTCTTCGGTGGGAAGGACACCGGATATTACACCTGCTCTT	1027
Db	992	TTCTGTCTTTGTCTATCATGAGTGGCAAGAACGACACCGGATATCTACACTGTGTCTCT	1051
Qy	1028	CCTCAAAAGCACCCAGCCAGCTACGCTTGGTGACCATCTAGAAAAAGGTTTATAAAG	1087
Db	1052	CTTCAAGCATCCAGTCAATCAGCTTTGGTTACCATCGTAGGNAAGGATTTATAATG	1111
Qy	1088	CTACCAGCTCGCAAGAGAGATGAAATTGACCCGTACGAAAGTTCTGCTCTCAGTCA	1147
Db	1112	CTACCAATTCAAGTGAAGATTATGAAATTTGACCAATATGAAGAGTTTGTGTTCTGTCA	1171
Qy	1148	GGTTTAAAGCGTACCCACGAAATCCGATGCAGCTGGATCTTCTCTCAAGCCTCATTTCCCTT	1207
Db	1172	GGTTTAAAGCCTACCCCAAAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCCCTT	1231
Qy	1208	GTGAACAGAGAGGCGTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAAGAACA	1267
Db	1232	GTGACAAAGGGTCTTGATAACCGATACAGCATATCCAAGTTTGGCATCAAGCAACC	1291
Qy	1268	AGCCAGGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTCACCAAATGTCA	1327
Db	1292	AGCCAGGAGATATATATTCATCGCAAAATGATGATGCCAATTTACCAAATGTCA	1351
Qy	1328	CGCTGAATATAAGAAAGAAACCTCAAGTGTACGAAATGCTCAGCAGCAGCGCTCCT	1387
Db	1352	CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGAAGAGCATCGGAAGTTCAGCGCTCT	1411
Qy	1388	GTTCCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAGTGTTCGGAAGAAATCTC	1447
Db	1412	GTTTCTCGGATGGATACCCATTACCATCTTTGGACCTGGAAGAGTGTTCAGACAGTCTC	1471
Qy	1448	CCAAATTCACGGAGGAAATCCAGAGAGGATTTGGAAATAAAGGCTPAACAGAAAAGTGT	1507
Db	1472	CCAACTGCACAGAGAGATCACAGAGGAGTCTGGAAATAGAAGGCTTAACAGAAAAGTGT	1531
Qy	1508	TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGCGCGGAAAGGCTCTCG	1567
Db	1532	TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGTTCCTCG	1591
Qy	1568	TCAAAATGCTGTGGTACAAATTTCTATGGGCAGCTCTTCGAAACCATCTTTTAAACTCAC	1627
Db	1592	TCAAGTCTGTGCATACAATTTCCCTTGGCACATCTTGTGAGACGATCCTTTTAAACTCTC	1651
Qy	1628	CAGGCCCTTCCCTTTTCATCCAAGACAACATCTCCTTCTATGGACCATTTGGGCTCTGTC	1687
Db	1652	CAGGCCCTTCCCTTTTCATCCAAGACAACATCTCAATCTATGCAACAAATTTGGTGTGTC	1711
Qy	1688	TCCCTTTCAATGTGTCTCATTTGTGATCTGCCCAAATACAAAAAGCAATTTAGGT	1747
Db	1712	TCTCTTCATTTGCTGTTTAAACCCTGCTAATTTGTCAACAGTACAAAAAGCAATTTAGGT	1771
Qy	1748	ACGAGTCAAGTCAAGATGATCCAGTGACTGCCCCCTCGATTAACGAGTACTTCTAGG	1807
Db	1772	ATGAAAGCCAGCTACAGATGGTACAGTGCACCGCTCCTCTCAGATAATGAGTACTTCTACG	1831
Qy	1808	TTGACTTCAGGACTATGAATATGACCTTTAAGTGGGAGTTCCCGAGAGACAATTAGAGT	1867
Db	1832	TTGATTTTCAGAGAAATGAATATGATCTCAATGGGAGTTTCCAAGAGAAATTTAGAGT	1891

Qy	1868	TTGGGAAGGTCTCTGGGGTCTGGGGCTTTTCGGGAGGTGATGAACGCCACCGCCTATGGCA	1927
Db	1892	TTGGGAAGGTACTAGGATCAGTGTCTTTTGGAAAGCTGATGAACGCAACAGCTTATGGAA	1951
Qy	1928	TTAGTAAACCGGAGTCTCAATTCAGTGTGGGTGAAGATGCTAAAGAGAAAGCTGACA	1987
Db	1952	TTAGCAAAACAGAGTCTCAATCCAGTGTGCCGTCAAAATGCTGAAGAAAGACGACACA	2011
Qy	1988	GCTGTGAAAGAAAGCTCTCATGCTCGAGGCTCAAAATGATGACCCACCTGGGACACCATG	2047
Db	2012	GCTCTGAAGACAGAGGCATCATGTCAGAACTCAAGATGATGACCCAGCTGGGAAGCCACG	2071
Qy	2048	ACAACTCGTAATCTGCTGGGGGCGATGACACTGTCAAGGCCAGTGTACTTGAATTTTGTG	2107
Db	2072	AGAAATGTGAACCTGCTGGGGGCGTGACACTGTCAAGGCCAAATTTACTTGAATTTTGTG	2131
Qy	2108	AAATATCTTCTATGTGTGACTCCTCAACTACCTAAGAAGTAAAGAGAGAAATTTTCACA	2167
Db	2132	AACTACTGTTGCTATGTGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAGAAATTTTCACA	2199
Qy	2168	GGACTGGACAGAGATTTTAAAGAAACATAAATTTTCAGTTCTTACCCCTACTTTCCAGGCAC	2227
Db	2192	GGACTGGACAGAGATTTTCAAGSAACACAAATTTTCAGTTTACCCCACTTTCCCAATCAC	2251
Qy	2228	ATTCAAATTCACGATGCTGTGTTCTCAGAGAAAGTTTCAGTTTACACCGCCGCTTGGATCAGC	2287
Db	2252	ATCCAAATTCAGCATGCTGTGTTCAAGAGAAGTTTCAGATACACCCGGAGCTCGGATCAAA	2311
Qy	2288	TCTCAGGGTTCAATGGGAATTCATTCATCTCAAGATGAGATTTGAATATGAAAACCCAGA	2347
Db	2312	TCTCAGGGCTTCATGGGAATTCATTCATCTCAAGATGAAATTTGAATATGAAAACCCAAA	2371
Qy	2348	AGAGCTGGCAGAAGAAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGACCTCCTTT	2407
Db	2372	AAAGCTG-----GAAGAGAGAGAGACTTGAATGTGCTTACATTTGAAGATCTCTTT	2425
Qy	2408	GCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGTGCCACA	2467
Db	2426	GCTTTGCAATATCAAGTTGCCAAAGGAATGAAATTTCTGGAATTTAAGTCGTGTGTTTACA	2485
Qy	2468	GAGACCTGGCAGCCAGGAATGTGTTGGTCACCCACCGGAAGTGGTGAAGATCTGTGACT	2527
Db	2486	GAGACCTGGCGCGCAGGAAGCTGCTGTCAACCCACCGGAAGTGGTGAAGATGTGTGACT	2545
Qy	2528	TTGACTTGGCCCGAGACATCTCAGGCGACTCCAGCTACGTCGTCAAGGGCAACGCACGCG	2587
Db	2546	TTGGATTTGGCTTCAGATATCATCAGTGATTCCAACTATGTTGTCAGGGCAATGCCCGTC	2605
Qy	2588	TGCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATCAAGAGTG	2647
Db	2606	TGCTGTGAAATGGATGGCCCCGAAAGCCTGTTTCAAGGCATCTACACCATTAAGAGTG	2665
Qy	2648	ACGTCGTGCTTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCCCTTACC	2707
Db	2666	ATGTCGTGCTATATGGAATATTAATCTGGGAAATCTTCACTTGGTGTGAATCTTACC	2725
Qy	2708	CTGGCATCTCTGTCCAGCGCTAACTTCTATAAAGTGAATCAGAGTGGATTTAAATGGAGC	2767
Db	2726	CTGGCATCTCGGTGTGCTTAATCTTCTCAAACTGATTCAAAATGATTTAAATGGATC	2785
Qy	2768	AGCAATCTATGCCACAGAGGATATATCTTGTAAATGCAATCCTGCTGGGCTTTTGACT	2827
Db	2786	AGCAATTTATGTATACAGAGAAATATACATTTAATGCAATCCTGCTGGGCTTTTGACT	2845
Qy	2828	CAAGGAGCGGCCATCTCTCCCAACCTGACTTCATTTTATAGATGTCAGCTGGCAGAGG	2887
Db	2846	CAAGGAACGCGCCATCTCTCCCTAAATTTGACTTTCGTTTTTAGGATGTCAGCTGGCAGATG	2905
Qy	2888	CAGAAGAAG-----CATGTATCAGAAACATCATCATCTACC	2924
Db	2906	CAGAAGAAGCGATGATCAGAAATGTGGATGGCGGTGTTTCGGGAATCTCTCACACCTACC	2965
Qy	2925	AAAACAGCGCGCCCTCAGCAGAGAGCGCGGCTCAGAGC-CAGTGCACACAGCGCCAGG	2983

QY 548 ACTTTAGGAAGATGGAAGAACAGGATGACATGCTCTGATCTCCGAGGGTGTCCGGAGC 607
DB 572 ACTTTAGGAAGATGGAAGAACAGGAGCCCTGGTCTGATATCTGAGAGCGTTCCAGAGC 631
QY 608 CCACTGTGGAGTGGGTCTCTGCACTGCCACAGGGAAGCTGTAAAGAAAGAGCCCTG 667
DB 632 CGATCGTGAATGGTGTTCGATTCACAGGGGAAGAGCTGTAAAGAAAGAGTCCAG 691
QY 668 CTGTTGTCAAGAGGAGGAAGTACTTTCATGATGTTGTCGGAACAGACATCAGATGCT 727
DB 692 CTGTTGTAAAGAGGAGGAAGTCTGCTCATGAATTAATTTGGGACGACATAGGTGCT 751
QY 728 GTGCTAGAAATGCACTGGCGCGGAATGCACCAAGCTGTTCCACCATAGATCTAAACAGG 787
DB 752 GTGCCAAGATGAAGTGGGAGGGAATGCACAGGCTGTTCCAAATAGATCTAAATCAA 811
QY 788 CTCCTCAGACACACTGCCCCAGTTATTCTCTGAAAGTGGGGAACCCCTTGTGGATCAGT 847
DB 812 CTCCTCAGACACACTGCCCAATTAATTTCTTAAAGTAGGGGAACCCCTTATGGATAAGT 871
QY 848 GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCC 907
DB 872 GCAAAAGCTGTTCATGTGAACCATGGATTCGGGCTCACCTGGGAATTAAGAAACAAAGCAC 931
QY 908 TGGAGGAGGCGAGTACTTTCAGATGATGATCTACTCCACAAACAGGACCATGATTCGGA 967
DB 932 TCGAGGAGGCGAAGTACTTTCAGATGATGATCTACTCCAAACAGAGACTATGATACGGA 991
QY 968 TTCTCTTGGGCTTTGTGCTTCCGTTGGGAAGAAACAGACACCGGATATTACACCTGCTCT 1027
DB 992 TTCTGTTTGTGTTATCATCAGTGGCAAGAACGACCGGATACACTACTACTGTTCTCT 1051
QY 1028 CTTCAAAGACCCAGCCAGTCAAGCTTGGTGGACCATCTCTAGAAAGAGGGTTTATAACG 1087
DB 1052 CTTCAAAGACATCCAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAATG 1111
QY 1088 CTACCAGCTCGRAGAGAGTATCAAAATGACCGTAGGAAAGTCTGCTTCTCAGTCA 1147
DB 1112 CTACCAATTCAGTGAAGATTAATAATGACCAATATGAAGATTTGTTTTCTGTGCA 1171
QY 1148 GGTTTAAAGGCTACCCAGCAATCCGATCGAGCTGGATCTTCTCAAGCCCTCATTTCCCT 1207
DB 1172 GGTTTAAAGCTACCCACAATCAGATGACGTGGACCTTCTCTCGAAATCAITTCCT 1231
QY 1208 GTAAACAGAGAGCCCTGGAGATGGGTACAGCATATCTAAATTTGGGATCATTAAGACA 1267
DB 1232 GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAGTTTGGCAATCATTAAGCAC 1291
QY 1268 AGCCAGGAGATACATTTCTATCAGAAATGATGAGCCCGCTTCCACCAAAATGTTCA 1327
DB 1292 AGCCAGGAGATATATATTCATCAGAAATGATGATGAGCCCAATTTACCAAAATGTTCA 1351
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTGAGCAAAATGCTCAGCCAGCGGCTGCT 1387
DB 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGCAGAGCATCGGCAAGTCAGGCTCT 1411
QY 1388 GTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGAGCAAAATCTC 1447
DB 1412 GTTCTCGGATGATACCCATTACCATCTTGGACCTGGAAGAGTGTTCAGACAAAGTCTC 1471
QY 1448 CCAATTCAGCGGAGGAATCCCAAGAGGATTTGGAATATAAGGCTTAACAGAAAGTGT 1507
DB 1472 CCAATTCAGAGAGATACAGAGAGGATGCTGGAATAGAAGGCTTAACAGAAAGTGT 1531
QY 1508 TTGGCCAGTGGTGTGAGAGTACTCTTAATATGATGAGCGCGGGAAGGGCTTCGG 1567
DB 1532 TTGGACAGTGGTGTGAGAGTACTCTAAACATGATGTAAGCCATATAAGAGGTTCTGG 1591
QY 1568 TCAATGCTGTGCTACAAATTTCTATGGGCGAGCTGTGCGAAACCATCTTTTAAACTCAC 1627
DB 1592 TCAAGTGTGTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCCTTTTAAACTCTC 1651

QY 1628 CAGGCCCTTCCCTTTCATCCAAAGACAAACATCTCCTTATCGACCATTTGGCTCTGTC 1687
DB 1652 CAGGCCCTTCCCTTTCATCCAAAGACAAACATCTCCTTATCGARCAATTTGGTGTGTC 1711
QY 1688 TCCCTTCATTTGTTCTTCTCATTTGTTGATCGCCACAAATACAAAAAGCAATTTAGGT 1747
DB 1712 TCCCTTCATTTGTTCTTTCATCCCTTAAATTTGTACAAAGTACAAAAAGCAATTTAGGT 1771
QY 1748 ACAGAGTCACTGTCAGATGATCCAGGTGACTGGCCCTGGATACGAGTACTTCTACG 1807
DB 1772 ATGAAAGCCAGTACAGATGGTACAGGTGACCCGCTCCATCAGATATGATGATCTTACG 1831
QY 1808 TTGACTTCAGGCACTATGAATATGACCTTAACTGGGAGTTCGCCAGAGAGAACTTAGAGT 1867
DB 1832 TTGATTCAGAGAAATATGAATATGATCTCAATGGGAGTTTCCAGAGAAATTTAGAGT 1891
QY 1868 TTGGAAAGTCTTGGGCTGTCGGGAGGTGATGAACGCCAGCCGCTATGGCA 1927
DB 1892 TTGGAAAGTACTAGGATCAGGTGCTTTTGGAAAAAGTGAACGCAACAGCTTATGGAA 1951
QY 1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAGAGAGAAAGCTGACA 1987
DB 1952 TTAGCAAAACAGGAGTCTCAATTCAGGTGGCGTCAAAATGCTGAAAGAAAAAGAGACA 2011
QY 1988 GCTGTGAAAAAGAGCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGACACATG 2047
DB 2012 GCTCTGAAGAGAGGCACTCATGTCAGAACTCAGATGATGATGATGATGATGATGATGAT 2071
QY 2048 ACAACATCTGTAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTGTACTTGAATTTTG 2107
DB 2072 AGAATATTGTGAACCTGCTGGGGGCTGCACACTGTCAGGACCAATTTACTTGAATTTG 2131
QY 2108 AATATTGTTGCTATGCTGACCTCCATCACTACTCTAAGAGTAAAGAGAGAGTTCACA 2167
DB 2132 AATACTGTGCTATGCTGATCTTCACTATCTAAGAGTAAAGAGAAAAATTTTACA 2191
QY 2168 GGACATGACAGAGATTTTAAAGAAACATAATTTCACTTCTTACCCCTACTTTCAGGAC 2227
DB 2192 GGACTTGGACAGAGATTTTCAAGGAACACAAATTTCACTTTTACCCCACTTTCATCAC 2251
QY 2228 ATTCAATTTCCAGCATGCTGTTCCAGGAAAGTTCAGTTACACCCGCTTTCGATCAGC 2287
DB 2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAAAGTTCAGATACACCCGCTTCGGATCAA 2311
QY 2288 TCTCAGGTTTCAATCGGAATTCATTTCAATCTCTGAAGTCAAGATTAATTAAGAAACAGA 2347
DB 2312 TCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGTGAATTAATTAAGAAACCAA 2371
QY 2348 AGAGCTGGCAGAAAGAGGAGGAAGATTTGAACTGCTGAGTTTGAAGACCTCTCTTT 2407
DB 2372 AAAGCTG-----GAAGAGAGGAGGACTTGAATGCTGTACATTTGAAGATCTTCTTT 2425
QY 2408 GCTTTCGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGTGTCTTCA 2467
DB 2426 GCTTTCATATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTCTACA 2485
QY 2468 GAGACCTGGCAGCCAGGAATGTGTGTCACCCAGGAGGAGTGTGAAGATCTGTGACT 2527
DB 2486 GAGACCTGGCCGAGGAACGCTGTTGCACCCAGGAAAGTGGTGAAGATATGTGACT 2545
QY 2528 TTGGAGTGGCCGAGACATCTCTGAGGACTCCAGCTACCTGTCAGGGCAACGACGCGC 2587
DB 2546 TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTACGGGGCAATGCCGTC 2605
QY 2588 TGCCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACAAATCAAGAGTG 2647
DB 2606 TGCTGTAAATGGATGGCCCGGAAAGCTGTTGAAGSCATCTACACATTAAGAGTG 2665
QY 2648 AGCTGTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
DB 2666 ATGCTGTGCTATATGGAATATTACTGTGGAAATCTTCTCACTTGGTGTGAATCTTACC 2725
QY 2708 CTGGCATTTCCCTCAGCGTAACTTCTATAACTGATTCAGAGTGGATTTAAATGGAG 2767

Db 2726 CTGCGATCCGGTTGATGCTTAACCTCTACAAAGTAATCAAAATGGATTTAAATGGATC 2785
Qy 2768 AGCCATTCTATGACACAGAGGATATACCTTTGTAATGCAATCTGCTGGGCTTTTGACT 2827
Db 2786 AGCCATTCTATGACACAGAGGATATACCTTTGTAATGCAATCTGCTGGGCTTTTGACT 2845
Qy 2828 CAAGGAAGCGCCATCTCTCCCAACCTGAGCTTATTTAGGATGTCAGCTGCAGAGG 2887
Db 2846 CAAGGAAGCGCCATCTCTCCCAACCTGAGCTTATTTAGGATGTCAGCTGCAGATG 2905
Qy 2888 CAGAAGAAG-----CATGTATCAGAAACATCCATCCATCTACC 2924
Db 2906 CAGAAGAAGCGATGATCAGAAATGGATGCGCTGTTCGGAATGCTCTCACACCTACC 2965
Qy 2925 AAAACAGCGCCCTCAGCAGAGAGGGGCTCAGAGC--CAGTCGCCACAGCGCCAGG 2983
Db 2966 AAAACAGCGACCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025
Qy 2984 T-GAAGATTACAGAGAAAGATTAGCGAGGAGGCTTTGGACCCCGCC---ACCCTAG 3038
Db 3026 TCGAAGATTCTGAGGAGCAATTTAGTTTAAAGACTTCATCCCTCACCTATCCCTTAA 3085
Qy 3039 CAGGCTGTAGACCGCAGAGCAAGATTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTCATCACTAAAGAAATCTATTATCAAC 3145
Qy 3097 CGTTGCTTCGCTGACTTTTCTCTAGATGCTGTGCGCACTACTC-----CAAAGTGA 3149
Db 3146 TGCTGCTTCACGACTTTTCTCTAGAGCGCTGCTGCTTTACTCTTTTCAAGAGGA 3205
Qy 3150 CTTCTATAAATCAAACTCTCTCGCACAGCGGGGAGGCAATATGAGACTTGTGG 3209
Db 3206 CTTTGTAAATCAAACTCTCTGTCACAGGAGGAGGCTGATAATGAACCTTTATTG 3265
Qy 3210 TGAGCGCCCTTACCTCGGGGCTTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269
Db 3266 GAGCATTGATCTGCATCAAGGCTTCTCAGCGGCTTGAGTGAATGTGTACTCTGAAG 3325
Qy 3270 TATGATATTTCTGTAATACGTTGAAGAAACAACCAACCCGTTTTCCTTAAGGAAGC 3329
Db 3326 TACAGTATTTCTGTAATACATAAAAACAAA-----AGCATTTTCTGAAGGAGAGC 3378
Qy 3330 TAAATATGATTTTAAATATCTGTTTAAATATCTATGTAACCTTTTCTATCTATTAG 3389
Db 3379 TAAATGATTTT---AAGTCTATGTTTAAATATATGTAATTTTTCAGCTATTAG 3436
Qy 3390 TGATATATTTATGATGGAATAAACTTTCTACTGTAAAAAATAAAAAAATAAAAAA 3449
Db 3437 TGATATATTTATGGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAAAAATAAAAAA 3496
Qy 3450 AAAA 3453
Db 3497 AAAA 3500

RESULT 14

ID AAT72117
AA72117 standard; cdNA; 3501 BP.

XX AC AAT72117;

XX 19-AUG-1997 (first entry)

XX Human flk-2 receptor coding sequence.

XX Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
XX pTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
XX haematopoietic hierarchy; extracellular domain; soluble form; ligand;
XX proliferation; differentiation; mammalian; haematopoietic stem cell;
XX macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 58..3039
FT /*tag= a
FT /product= Human_flk-2_receptor
FT 58..138
FT /*tag= b
FT 139..3036
FT mat_peptide
FT /*tag= c
XX US5621090-A.
XX 15-APR-1997.
XX 02-APR-1991; 91US-0679666.
XX 26-JUN-1992; 92US-0906397.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
XX (UYPR-) UNIV PRINCETON.
PA Lemischka IR;
XX WPI; 1997-235228/21.
DR P-PSDB; AAW19873.
XX Protein containing the extracellular domain of human flk-2 - used
PT for identification of primitive haematopoietic cell proliferation
PT and differentiation stimulatory ligands, e.g. for treating anaemia
XX Claim 1; Fig 1B; 55pp; English.
XX This sequence encodes the human fetal liver kinase 2 (flk2). flk-2 is
CC a receptor protein tyrosine kinase (ptk) and is important in transducing
CC putative self-renewal signals from the environment. flk-2 is expressed
CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow.
CC and it is thought that flk-2 is expressed in the entire primitive portion
CC of the haematopoietic hierarchy. The invention concerns a recombinant
CC nucleic acid, preferably mRNA, which encodes a protein containing only
CC the extracellular domain of human flk-2 and lacking the flk-2 intra-
CC cellular catalytic domain. The resultant protein represents a soluble
CC form of flk-2 which is used to isolate specific ligands for flk-2. These
CC ligands can be used to stimulate proliferation and/or differentiation of
CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused
CC by cancer treatment or radiation.
XX SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 64.5%; Score 2227.2; DB 18; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy 8 GGTACCGCGCTCGGAGGCGCATCGCGGCGTGGCGAGCGACGCGCGGCTGC 67
Db 35 GGGGACCCCGGCTCGGAGGCGCATCGCGGCGTGGC---GCGGACGCGGCGCGCTGC 91
Qy 68 TGCTGCTTGTGTTTGTGTCAGTAATGATTTTGTAGACCGTTACAAACCAAGACCTGCCTG 127
Db 92 CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG 151
Qy 128 TGATCAAGTGTCTTTTAAATCATGATGAGAACAAATGGCTCATCGCGGGAAGCCATCAT 187
Db 152 TGATCAAGTGTCTTTTAAATCATGATGAGAACAAATGATTCATCAGTGGGGAAGTCATCAT 211
Qy 188 CCGTACGAATGTTGGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGGAGCGCCAGGTG 247
Db 212 CATATCCCATGGTATCAGAAATCCCGAAGACCTCGGCTGTGCGTTTGAGACCCAGAGCT 271
Qy 248 AAGGACGCTATATGAGGCGGCGCACCGTGGAGGTGCGCGAGTCTGCGTCCATCACCTGC 307

Db 272 CAGGACAGTGTACGAAGTGCCTGCGTGGAAAGTGATGATCTGCTCCATCACACTGC 331
QY 308 AAGTCCAGCTCGCCACCCAGGGGACCTTTCCTGCTCTGGGTCTTTAAGACAGCTCCC 367
Db 332 AAGTCTCGGTGATGCCCCAGGGACATTTCTGCTCTGGGTCTTTAAGACAGCTCCC 391
QY 368 TGGGCTGCCAGCCGACCTTTGATTTACAAAACAGAGGAATCCTTCCATGGCCATCTTGA 427
Db 392 TGAATTGCCAGCCACATTTGATTTACAAAACAGAGGAGTGTTCATGGTCAATTTGA 451
QY 428 ACGTGACAGACCCAGGAGAGAAATACCTACTCTATTCATATTCAGAGCGCCGCAACT 487
Db 452 AAATGACAGAAACCCAGCTGAGAAATACCTACTCTTTTATTACAGAGTGAAGCTACCAAT 511
QY 488 ACACAGTACTGTTACAGTGAATTAAGAGATACACAGCTGTATGTCTAAGGACCTT 547
Db 512 ACACAAATATTGTTTACAGTGAATTAAGAAATACCTCTGTTTACACATTAAGAAGACCTT 571
QY 548 ACTTTAGGAAGATGAAACACAGGATGCACCTGCTGCTATCCGAGGCTGTTCCGGAGC 607
Db 572 ACTTTAGAAAATGGAACACAGGAGCCCTGCTGCTATCTCAGAGCGTTCAGAGC 631
QY 608 CCACTGTGGAGTGGTGTCTGCACTCCACAGGGAAGCTGTAAAGAAAGAGCCCTG 667
Db 632 CGATCGTGAATGGTGTCTTGGCATTCACAGGGGGAAGCTGTAAAGAAAGTCCAG 691
QY 668 CTGTTGTCAGAAAGGAGGAAAAGGTACTTCATGAGTTGTTCCGACACACATCAGATGCT 727
Db 692 CTGTTGTTTAAAGAGGAGAAAAGTCTCATGAATTTATTTGGGACGACATAAGGTGCT 751
QY 728 GTGCTAGAAATGCACTGGCGCGGAATGCACCAAGCTGTTCAACATAGATCTAAACACAGG 787
Db 752 GTGCAGAAATGAAGTGGCGAGGAATGCACAGGCTGTTCAAAATAGATCTAAATCAAA 811
QY 788 CTCCTCAGAGACACTGCCCGAGTTATTCCTGAAAGTGGGGAAACCTTTGGGATCAGGT 847
Db 812 CTCCTCAGACACATTCGCCAAATTTATTTCTTAAGTAGGGGAACCTTTATGGAATAGGT 871
QY 848 GTAAGGCCATCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGCAAGGCC 907
Db 872 GCAAGCTGTTTATGTGAACCATGATTCGGGCTCACCTGGGAGTTAGAAACAAAGCAC 931
QY 908 TGGAGGGGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATGATTCGGA 967
Db 932 TCGAGGAGGGCACTACTTTGAGATGAGTACTTATCAACAAACAGAACTATGATACGGA 991
QY 968 TTCTCTTGGCCCTTTGCTCTCCGTGGGAAGAACGACACCGGATATTACACCTGCTCTT 1027
Db 992 TTCTGTTTGTCTATCATCAGTGGCAAGAAACGACACCGGATACTACACTTGTCTCT 1051
QY 1028 CTCNAACACCCAGCCAGCTCAGCGTTGGTGACCATCTAGAAAAGGTTTATAAGC 1087
Db 1052 CTTCAAGCATCCCCAGTCAATCAGCTTTGGTTACCATTGTTAGGAAGGATTTATAAATG 1111
QY 1088 CTACAGCTCGCAAGAAGATGATAAATGACCCGTACGAAAGTTCTGCTTCTCAGTCA 1147
Db 1112 CTACCAATTCAAGTGAAGATTAAGAAATGACCAATATGAGAGTTTGTCTTCTGTC 1171
QY 1148 GGTTTAAAGCGTACCAGCAATCCGATCGATCGATGGATCTCTCTCAAGCCCTCAATTCCTT 1207
Db 1172 GGTTTAAAGCCTACCCACAAATCAGATGATAGTGGACCTCTCTCGAAAATCAATTCCTT 1231
QY 1208 GTGAACAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCATAGACA 1267
Db 1232 GTGAGCAAAAGGCTTGTGATACGGATACAGCATATCCAAGTTTGTGCAATATAAGCACC 1291
QY 1268 AGCCAGGAGACTACATATCTATGCAGAAAATGATGAGCCCGAGTTCAACCAAAATGTTCA 1327
Db 1292 AGCCAGGAGATATATATTCATGCAGAAAATGATGATGCCCAATTTACCAAAATGTTCA 1351
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTTCAGCCAGCCAGGCGCT 1387
Db 1351 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTTCAGCCAGCCAGGCGCT 1387

Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTGCAAGAGCATCGCGAAGTCAAGCGCTCT 1411
QY 1388 GTTCTCTCATGGCTTACCCTTACCCTCTTGGACCTGGAGAGTGTTCGGAACAATCTC 1447
Db 1412 GTTCTCTCGATGGATACCCATACCCTTGGACCTTGGAGAGTGTTCAGACAAGTCTC 1471
QY 1448 CCAATTGTCACGGAGGAAATCCCAAGAGAGTGTGGAAATAAAAAGGCTTAACAGAAAAGTGT 1507
Db 1472 CCAACTGCACAGAGATCAGAGAGAGTGTGGAATAGAAAGGCTTAACAGAAAAGTGT 1531
QY 1508 TTGGCCAGTGGGTGTCGAGCAGTACTTAATATAGTGTAGGCGGGAAGGCTTCTGG 1567
Db 1532 TTGGCAGTGGGTGTCGAGCAGTACTTAAACATGAGTGAAGCATAAAAAGGCTTCTGG 1591
QY 1568 TCAATGCTGCTGCTACAAATTTATGGCAGCTTCTTGGAAACCAATCTTTTAAACCTCAC 1627
Db 1592 TCAAGTGTGTGCATACAATTTCCCTTGGCAGATCTTGTGAGAGATCCCTTTTAAACCTC 1651
QY 1628 CAGGCCCCCTTCCCTTTCATCCAAAGACACATCTCTTCTATGCGACATTTGGGCTCTGTC 1687
Db 1652 CAGGCCCCCTTCCCTTTCATCCAAAGACACATCTCTTCTATGCAACAATTTGGTGTGTC 1711
QY 1688 TCCCTCTCATGTTGTTCTCATTTGTTGATGTGCGACAAATACAAAACAATTTAGGT 1747
Db 1712 TCCCTCTCATGTTGTTTAAACCTGCTAAATTTGTCACAAGTACAAAAGCAATTTAGGT 1771
QY 1748 ACGAGAGTCACTGCAGATGATCCAGGTCACTGGCCCCCTGGATACAGAGTACTCTACG 1807
Db 1772 ATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCAGATATAGTACTCTTACG 1831
QY 1808 TTGACTTCAAGGACTATGAATATGACCTTAAAGTGGGAGTTCGCGAGAGAACTTAGAGT 1867
Db 1832 TTGATTTCAAGAAATATGAATATGATCTCAATGGGAGTTTCCAAAGAAAATTTAGAGT 1891
QY 1868 TTGGAAAGTCTCGGGCTTGGGAGGTGATGAACGCCAGCCCTATGGCA 1927
Db 1892 TTGGAAAGTCTAGGATCAGGTGCTTTTGGAAAAGTGAACGCCAACAGCTTATGGAA 1951
QY 1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAAAGCTGACA 1987
Db 1952 TTAGCAAAACAGAGTCTCAATCCAGGTTGCGGCTCAAAATGCTGAAAGAAAAGCAGACA 2011
QY 1988 GCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
Db 2012 GCTCTGAAGAGAGGCACTCATGTCAGAACTCAGATGATGACCCAGCTGGGAGCCAG 2071
QY 2048 ACAACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGCGCAGTGTACTTGTATTTG 2107
Db 2072 AGCAATATGCGAACCTGCTGGGGCGTGCACACTGTGAGGACCAATTTACTTGTATTTG 2131
QY 2108 AATATGTTGCTATGTTGACCTCTCACTACCTAGAGATGAAAGAGAGAGTTCACACA 2167
Db 2132 AATAGTGTGCTATGTTGATCTTCACTATCTAAGAAAGTAAAGAGAAAATTTTCA 2191
QY 2168 GCACATGACAGAGATTTTAAAGGAACATAATTTCACTTCTTACCCTACTTTTCCAGCAC 2227
Db 2192 GCACATGACAGAGATTTTCAAGGAACACATAATTTCACTTTTACCCCACTTTTCCAATCAC 2251
QY 2228 ATTCAAAATCCAGCATGCTGTTTACAGAAAGTTCAGTTACACCCGCTTGGATCAGC 2287
Db 2252 ATCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGCTCGGATCAA 2311
QY 2288 TCTCAGGTTCAATGGGAATTTCAATTCATCTGAGATGAGATGATATGAAAACACAGA 2347
Db 2312 TCTCAGGCTTCAATGGGAATTTCAATTCATCTGAGATGAAATTTGAATATGAAAACCAA 2371
QY 2348 ACAGCTGTCAGAAAGAGGAGGAGATTTCGAACGCTGCTGAGTGTGAGAGCTTCTCTT 2407
Db 2372 AAAGCTG-----GAAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTT 2425
QY 2408 GCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTGTGTGTCACA 2467
Db 2426 GCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAGTGTGTGTCACA 2485

```
QY 2468 GAGACCTGGCAGCCAGGAATGTGTTGCTACCCACCGGAAGGTGGTGAAGATCTGTGACT 2527
Db 2486 GAGACCTGGCCGCGCAGACGCTGTGTCTACCCACCGGAAGGTGGTGAAGATATGTGACT 2545
QY 2528 TTGGAGTGGCCGCGAGACATCTCTGAGGAGCTCCAGCTACGTCGTGAGGGCAACGCGACGGC 2587
Db 2546 TTGGATTGGCTCGAGATATCATGAGTGAATTCCTCAACTATGTGTGACGGGGCAATGCCGCTC 2605
QY 2588 TCCGGGTGAAGTGGGACCGGAGAGCTTATTTGAAGGGATCTACAAATCAAGAGTG 2647
Db 2606 TGCCTGTAAATGATGGCCCGGAAAGCCTGTTGAAGGCATCTACACCAATTAAGAGTG 2665
QY 2648 AGCTCTGGTCTTACGACCTCTCTCTGGAGATATTTTCACTGGGTGTGAACCCCTTACC 2707
Db 2666 ATGCTGTGTCATATGAATATTAATCTGTGGAAATCTTCTCACTTGTGTGAATCCTTACC 2725
QY 2708 CTGGCATTCTCTGCGAGCTAACTCTATATAACTGAATCAGAGTGGATTTAAATGGAGC 2767
Db 2726 CTGGCATTCTCGGTGATGCTAACTCTACAACTGATTTCAAAATGGATTTAAATGGATC 2785
QY 2768 AGCCATTCTATGCCAGAGGATATATTTGTAATGCAATCTCTGGGCTTTTGTGACT 2827
Db 2786 AGCCATTCTATGCTACAGAGAAATATACATTAATGCAATCTCTGGGCTTTTGTGACT 2845
QY 2828 CAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTGGCAGAGG 2887
Db 2846 CAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTGGCAGAGG 2905
QY 2888 CAGAAGAAG-----CATGTATCAGAACTCCATCCATCATTACC 2924
Db 2906 CAGAAGAAGCGATGTATCAGATGTGGATGGCGCTTTCGGAATCTCTCTCACACCTACC 2965
QY 2925 AAAACAGGGCCCTTCACAGAGAGCGGGCTCAGC-CCAGTCCGCCACGCGCCAGG 2983
Db 2966 AAAACAGGGCCCTTCACAGAGAGATGGATTTGGGGTACTCTCTCCGCGAGGCTCAGG 3025
QY 2984 T-GAAGATTACAGAGAGAGAGTTAGCGAGGAGCGCTTGGACCCGCGC-----ACCCTAG 3038
Db 3026 TCGAAGATTCTGAGAGAGACATTTAGTTTAAAGGACTTCACTCCCTCCACCTATCCCTAA 3085
QY 3039 CAGGCTGTAGCGCAGAGCCAGATAGCTCTCGCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGATTACCAAAACAAGATTAATTTTCATCACTAAAGAAATCTATTATCAAC 3145
QY 3097 CGTTGCTTCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA 3149
Db 3146 TGTGCTTCTCACCAGACTTTTCTCTAGAAGCGCTCTGCTTACTCTGTGTTTCAAGGGA 3205
QY 3150 CTCTATAAAATCAAACTCTCTCTCGCAGCGGGGAGAGCCCAATATGAGACTTGTGG 3209
Db 3206 CTCTATAAAATCAAACTCTCTCTCGCAGCGAGGAGCTGATAATGAATTTATTG 3265
QY 3210 TGAGCCCGCTACCCCTGGGGCTTTTCCAGAGCTTGGAGGAAAGCAATGATCTGAAA 3269
Db 3266 GAGCATTGATCTGCATCCCAAGGCTTCTCAGGCGGCTTGAGTGAATTTGTTACCTGAAG 3325
QY 3270 TATAGTATATTCTGTAAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGAAAGC 3329
Db 3326 TACAGTATATTCTGTAAATACATAAACA-----AGCATTTTGTGAAGGAAAGC 3378
QY 3330 TAAATATGATTTTAAATATCTATGTTTAAATATCTATTAATCTTTCATCTATTATG 3389
Db 3379 TAATATGATTTT--AGTCTATGTTTAAATATATATGTAATTTTTCAGCTATTATG 3436
QY 3390 TGATATATTTATGATGAAATATAAATTTTCTACTGTAAAAAATAAATAAATAAATAA 3449
Db 3437 TGATATATTTATGCTGGAAATAAATTTTCTACTACAGAAAAAATAAATAAATAAATAA 3496
QY 3450 AAAA 3453
Db 3497 AAAA 3500
```

RESULT 15

```
AAAX77515
ID AAX77515 standard; cDNA; 3501 BP.
XX
AC AAX77515;
XX
DT 05-AUG-1999 (first entry)
XX
DE Human flk-2 cDNA.
XX
KW Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
monoclonal; polyclonal; antibody; tyrosine kinase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 58..3039
FT /*tag= a
FT /product= "flk-2"
XX
PN US912133-A.
XX
PD 15-JUN-1999.
XX
PF 10-FEB-1998; 98US-0021324.
XX
PR 19-NOV-1992; 92US-0977451.
PR 02-APR-1991; 91US-0879666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252498.
PR 15-FEB-1996; 96US-0601891.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Lemischka IR;
XX
DR WPI; 1999-357194/30.
DR P-PSDB; AAY08617.
XX
XX Isolating hematopoietic cells expressing fetal liver kinase 1
PT receptors
XX
PS Disclosure; Fig 1b; 59pp; English.
XX
CC This invention describes a novel method of isolating cells expressing
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
CC binding the cells to a polyclonal or monoclonal antibody specific to
CC the Flk-1 receptor and isolating the cells that have bound to the
CC antibody. The method can be used to isolate hematopoietic stem cells in
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
CC the invention belong to the receptor protein family. This sequence
CC encodes the human flk-2 protein which is used in the method of the
CC invention.
XX
SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
Query Match 64.5%; Score 2227.2; DB 20; Length 3501;
Best Local Similarity 79.68; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 8 GGCTACCGCGCGCTCCGGAGGCCATCGGCGGTGGCGCAGCGCAGCGCGGGTGC 67
Db 35 GGGGACCCCGGCTCCGGAGGCCATCGGCGGTGGCGCGCGCGCGCGCGTGC 91
QY 68 TGCTGCTGTGTTTGTTCAGTAAATGATCTTGAGACCGGTGTACAAACAGACCTGCTG 127
Db 92 CGCTGCTGCTGTTTGTTCAGTAAATGATCTTGAGACCGGTGTACAAACAGATCTGCTG 151
```

Qy	128	TGATCAAGTGTGTTTAAATCATGCTATGAGAACAAATGGCTCATCAGCGGGGAAAAGCCATCAT	187
Db	152	TGATCAAGTGTGTTTAAATCAATCATAGAACAATGATTCATCAGTGGGGAAGTCATCAT	211
Qy	188	CGTACCGAATGSGCGAGGATCCCAAGAGACCTCCAGTGTACCCGAGCGGCCAGAGTG	247
Db	212	CATATCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTGGTGTAGACCCAGAGCT	271
Qy	248	AAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGSTCCATCACCCCTGC	307
Db	272	CAGGGACAGTGTACGAAGCTGCCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC	331
Qy	308	AAGTGCAGCTCGCCACCCAGGGGACCTTTCCTGCCCTCTGGGTCTTTAAGCAGAGTCCC	367
Db	332	AAGTGTGGTTCGATGCCCGAGGAAATTCCTGTCTCTGGGGTCTTTAAGCAGAGTCCC	391
Qy	368	TGGGCTGCCAGCGCGACTTTTGATTACAAACAGAGGAATCGTTTCCATGCCCATCTTGA	427
Db	392	TGAATTGCCAGCACATTTTGATTACAAACAGAGGAGTGTGTTCCATGGTCATTTGA	451
Qy	428	ACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGGAGAACGGCCCAACT	487
Db	452	AAATGACAGAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT	511
Qy	488	ACACAGTACTGTTCAACAGTCAATGTAGAGATACACAGCTGTATGTGCTTAAGGAGACCTT	547
Db	512	ACACAATATTGTTACAGTGAATAGAAATACCTGCTTTACACATTAAGAAGACCTT	571
Qy	548	ACTTTAGAAAGATGGAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGTCCGGAGC	607
Db	572	ACTTTAGAAAAATGGAANAACAGGACGCCCTGCTGTCATATCTGAGAGCGTTCCAGAGC	631
Qy	608	CCACTGTGGAGTGGGTGCTCTGCAGCTCCCAAGGGAAAGCTGTAAAGAAAGAGGCGCTG	667
Db	632	CGATCGTGGAAATGGGTGCTTTGCGATTCAACAGGGGAAAGCTGTAAAGAAAGATCCAG	691
Qy	668	CTGTTGTCAGAAAGGAGGAAAAGGTACTTCATGATGTGTTTCGGNACAGACATCAGATGCT	727
Db	692	CTGTTGTATAAAGAGGAAAAAGTGCTTCATGAATTAATTGGGACGGACATAAAGTGCT	751
Qy	728	GTGCTAGAAATGCACTGGGCGCGGAATGCAACAGCTGTTCACCATAGATCTAAACCCAGG	787
Db	752	GTGCCAGAAATGAATGGGCGAGGAATGCACAGGCTGTTCACAATAGATCTAAATCAA	811
Qy	788	CTCCTCAGACACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCCCTTGTGATCAGGT	847
Db	812	CTCCTCAGACACATTGCCCAATATTTCTTAAAGTAGGGGAACCCCTTATGGATAAGGT	871
Qy	848	GTAAAGCATCCATGTCAACCATGGATTTCGGGCTCACTGGAGCTGGAAGACAAAGCCC	907
Db	872	GCAAGCTGTTATGTGAACCATGGATTTCGGGCTCACCTGGGAATTAGAAACAAAGCAC	931
Qy	908	TGGAGGAGGCGACTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATTATTCGGA	967
Db	932	TCGAGGAGGGCAACTACTTTGAGATGATACCTATTCAACAAACAGAACTATGATACGGA	991
Qy	968	TTCTCTTGGCCCTTGTGCTTTCGTGGGAGGAACGACACGGGATATACACCTGCTCTT	1027
Db	992	TTCTGTTTGTCTTTTGTATCATGATGCAAGAACGACACGGGATACACATTTGTTCCCT	1051
Qy	1028	CCTCAAGCACCCGACGCTACGCTGGTGACCATCTTAGAAAAAGGGTTTATAACG	1087
Db	1052	CTTCAAAAGCATCCAGTCAATCAGCTTTGGTTACCATCTGAGGAAGGGATTATAATG	1111
Qy	1088	CTACAGCTCGCAAGAGATATGAAATTTGACCCGTACGAAAGAAATCTGCTTCTCAGTCA	1147
Db	1112	CTACCAATTCGAAGTGAAGATTATGAAATTTGACCAATATGAAGAGTTTTGTTTTCTGTCA	1171
Qy	1148	GGTTTAAAGCGTACCCAGGAATCCGATGACGCTGGATCTTCTCTCAAGCCCTCATTTCCCT	1207
Db	1172	GGTTTAAAGCGTACCCAGCAATCAGATGTAGTGGACCTCTCTCTCGAAAATCATTTCCCT	1231

Qy	1208	GTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGCATCAATAAGACA	126
Dp	1232	GTGAGCAAAAAGGGTGTTGATAAAGCGATACAGCATATCCAAGTTTTCGAATCAAGAAGCACC	1291
Qy	1268	AGCCAGGAGAGTAGATATCTATGTCAGAAAAATGATGACGCCAGTGTCCACCAAATGTTCA	1327
Dp	1292	AGCCAGGAGAAATATATATTCCATGCGAGAAATGATGATGCCCCRAATTTACCAAAATGTTCA	1351
Qy	1328	CGCTGAATATAAGAAAAGAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCACAGCGCTCCT	1387
Dp	1352	CGCTGAATATAAGAAAAGAACCTCAAGTGTGCGAGAAAGCATCGGCAAGTCAAGCGCTCCT	1411
Qy	1388	GTTTCCCTGATGCGTACCCGCTACCTCTTTGGACTGCGGAGAGAGTGTTCGACAAAATCTC	1447
Dp	1412	GTTTCTCGGATGATACCCATTAATACCTTTGGACCTGGAAAGAGTGTTCAGACAAGTCTC	1471
Qy	1448	CCAATTCACGGAGGAAATCCCACAGAGGAGTTTGGAAATAAAGAGCTTAACAGAAAAGTGT	1507
Dp	1472	CRAACTGCACAGAGAGATCACAGAGGAGTCTGGAAATAGAAAGGCTACAGAAAAGTGT	1531
Qy	1508	TTGCCCAAGTGGGTGTCGAGCAGTACTCTAAATATAGTGTAGGCGGGAAGGGCTTCCTGG	1567
Dp	1532	TTGACAGTGGGTGTCGAGCAGTACTCTAAACATGATGAAGCCATAAAGGGTTCCTGG	1591
Qy	1568	TCAAATCTGTGCGTACAAATCTATGGCGAGCTTCTCGGAAACCATCTTTTTAAACTCAC	1627
Dp	1592	TCAAGTCTGTGCATACAATTTCCCTTGGCACATCTTGTGACAGCATCTTTTTAAACTCTC	1651
Qy	1628	CAGGCCCTTCCTTTTCATCCAAGACAACATCTCTTCTATGCGACCATTTGGGCTCTGTC	1687
Dp	1652	CAGGCCCTTCCCTTTTCATCCAAGACAACATCTCATTTCTATGCAACAATTTGGTGTGTC	1711
Qy	1688	TCCCTTCATTGTGTCTCATTTGTGTGATCTGCCACAAATACAAAAGCAATTTAGT	1747
Dp	1712	TCCTCTCATTTGCTTTTAACTGCTAATTTGCTCACAAGTACAAAAGCAATTTAGT	1771
Qy	1748	ACGAGTCAAGTGCAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTACTTCTACG	1807
Dp	1772	ATGAAAGCCAGCTACAGTGGTACAGGTACCGGCTCCTCAGATAATAGTACTTCTACG	1831
Qy	1808	TTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAACTTAGAGT	1867
Dp	1832	TTGATTTTCAGAGAAATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAAAATTTAGAGT	1891
Qy	1868	TTGGGAAGGTCTGGGCTGTGGCGCTTTCCGGAGGGTGATGAACGCCACAGCCCTTAGGCA	1927
Dp	1892	TTGGCAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCCACAGCTTATGGAA	1951
Qy	1928	TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAAAGCTGACA	1987
Dp	1952	TTAGCAAAACAGGAGTCTCAATCCAGTTGCGGTCAAAATGCTGAAAGAAAAAGCAGACA	2011
Qy	1988	GCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG	2047
Dp	2012	GCTCTGAAAGAGAGGCACTCATGTCGAACCTCAAGATGATGACCCAGCTGGGAGAGCCACG	2071
Qy	2048	ACAACATCTGTAATCTGCTGGGGCATGCACTGTCTAGGCGCCAGTGTACTTGATTTTGTG	2107
Dp	2072	AGAAATATTGGAACCTGCTGGGGCGGTGCACACTGTCAGGACCAATTTACTTTGATTTTG	2131
Qy	2108	AATATTGTGCTATGGTGACCTCTCAACTACCTACCTAAGAAAGTAAAGAGAGAAAGTTTCACA	2167
Dp	2132	AATACTGTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAAAATTTTCACA	2191
Qy	2168	GGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCTACTTTCAGGCAC	2227
Dp	2192	GGACTTGGACAGAGATTTTCAAGGAACACAATTTTCAGTTTATACCCCACTTTTCCAATCAC	2251
Qy	2228	ATTCAAATTCAGCATGCGTGGTTTCAGGAGAGTTTCAGTTTACACCCCGCTTGGATCAGC	2287
Dp	2252	ATCCAAATTCAGCATGCGTGGTTTCAGAGAGAGTTTCAGTACACCCCGGCACTTCGAGTCAAA	2311
Qy	2288	TCTCAGGGTTCAAATGGGAATTTCAATTTCAATCTGAAGATGAGATTTGAATATGAAACCCAGA	2347

```
Db 2312 TCTCAGGCTTCATGGAATTCATTTCACTCTCAAGATGAAATTAATGAATGAAACCCAAA 2371
Qy 2348 AGAGGCTGCAGAGAGAGAGAGAGATTTGACGTGCTGACGTTTGAAGACCTCCTTT 2407
Db 2372 AAAGCTG-----GAAGAGAGAGAGAGCTTGAATGCTGTACATTTGAAGATCTTCTTT 2425
Qy 2408 GCTTTCCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGATCGTGTGTCACA 2467
Db 2426 GCTTTGCATATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2485
Qy 2468 GAGACCTGCAGCCAGGAATGTGTTGGTCAACCACGGGAAGGTGGTGAAGATCTGTGACT 2527
Db 2486 GAGACCTGGCCGCGCAGGACGCTGCTGTCAACCACGGGAAGGTGGTGAAGATATGTGACT 2545
Qy 2528 TTGGACTGGCCGAGACATCTCTGAGCGACTCCAGCTACGTCAGGGGCAAGCCGACGGC 2587
Db 2546 TTGGATTTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTACGGGGCAATGCCCGTC 2605
Qy 2588 TGCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTCAAGGGATCTACAAATCAAGAGTG 2647
Db 2606 TGCTGTAAATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCATTTAAGAGTG 2665
Qy 2648 ACCTCTGCTPCTACGSCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAACCCCTTACC 2707
Db 2666 ATGCTGTGTCATATGGAATATTACTGTGGGAAATCTCTCACTTGGTGTGAATCCTTACC 2725
Qy 2708 CTGGCATCTCTGCGACGCTAATCTCTATTAAGTGAATCAGAGTGAATTAAGTGAAGC 2767
Db 2726 CTGGCATCTCGGTTGATGTAATCTCTACAAATGATTCAAATGGATTTAAATGGATC 2785
Qy 2768 AGCCATCTCTATGCCAGAGGATATCTTTGTAATGCAATCCTGCTGGGCTTTTGGACT 2827
Db 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGAATCTCTGCTGGGCTTTTGGACT 2845
Qy 2828 CAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGTGGCAGAGG 2887
Db 2846 CAAGGAACCGGCATCTTCCCTAATTTGACTTCGTTTTTAGGATGTCAGTGGCAGAGT 2905
Qy 2888 CAGAAGAG-----CATGTATCAGAACATCCATCATCTACC 2924
Db 2906 CAGAAGAGCGATGATACAGATGTGGATGGCGCTGTTTCGGAATGTCTCTCACACCTACC 2965
Qy 2925 AAAACAGCGCGCCCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983
Db 2965 AAAACAGCGGACCTTTCACAGAGAGATGGAATTTGGGGCTACTCTCTCCGCGGCTCAGG 3025
Qy 2984 T-GAAGATTCACAGAGAAAGATTTAGCGAGGAGGCTTTGGACCCCGCC-----ACCTAG 3038
Db 3026 TCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
Qy 3039 CAGGCTGTAGCGGAGAGCCCAAGATTTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGATTACCAAAACAAGATTAATTTTCATCACTAAAAAATAATCTATTATCAAC 3145
Qy 3097 CGTTGCTTCGCTGGACTTTTCTAGATGCTGTCTGCCATTACTC-----CAAAAGTGA 3149
Db 3146 TCGTCTTCCACGAGCTTTCTCTAGAGCCGCTCGGTTTACTCTTTTTCAAAGGGA 3205
Qy 3150 CTCTATAAAATCAAACTCTCTCTCGCACAGCGGGGAGGCCAATAATGAGACTTGTGG 3209
Db 3206 CTTTGTAAATCAAAATCATCTGTCAAGGCGAGGAGGCTGATTAATCAACTTTATG 3265
Qy 3210 TGAGCCCGCCTACCCCTGGGGGCTTTTCCAGAGCTTTGAGGGGAAAGCCATGTATCTGAAA 3269
Db 3266 GAGCATTTGATCTGCATCCCAAGGCTTCTCAGGCGGCTTGAGTGAATTTGTGTACCTGAAG 3325
Qy 3270 TATAGTATATTCTTTAAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGGAAGC 3329
Db 3326 TACAGTATATTCTTTGTAATACATAAAACAAA-----AGCATTTTGTAAAGGAAGC 3378
Qy 3330 TAAATATGATTTTAAAAATCTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
```

```
Db 3379 TAATATGATTTTTT--AAGTCTATGTTTTAAAAATAATATATGTAATTTTTTCAGCTATTAG 3436
Qy 3390 TGATATATTTTATGGATGGAATAAACTTTTCTACTGTAAAAAATAAATAAATAAATAA 3449
Db 3437 TGATATATTTTATGGGTGGGATAAAAATTTCTACTACAGAAAAAATAAATAAATAA 3496
Qy 3450 AAAA 3453
Db 3497 AAAA 3500
```

Search completed: May 25, 2003, 12:14:57
Job time : 705.168 secs

2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
2101 ATTTTGAATATGTTGCTATGCTGATGCTGCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
2101 ATTTTGAATATGTTGCTATGCTGATGCTGCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
2161 TTTCACAGACATGACAGAGATTTTAAAGAACATAATTTTACCTTCTTACCTACTTTTC 2220
2161 TTTCACAGACATGACAGAGATTTTAAAGAACATAATTTTACCTTCTTACCTACTTTTC 2220
2221 CAGGCACATTTCAATTTCCAGCATGCTGCTGCTCAGAGAGTTCAGTTACACCCGCCCTTG 2280
2221 CAGGCACATTTCAATTTCCAGCATGCTGCTGCTCAGAGAGTTCAGTTACACCCGCCCTTG 2280
2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTCTGAAAGTGAAGTGAATGAATGA 2340
2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTCTGAAAGTGAAGTGAATGAATGA 2340
2341 RACCAAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2341 RACCAAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2401 CTCTCTTCTGCTTACCAAGTGCCTCAAGGATGGAATTCCTGGAGTCAAGTCTGT 2460
2401 CTCTCTTCTGCTTACCAAGTGCCTCAAGGATGGAATTCCTGGAGTCAAGTCTGT 2460
2461 GTCCACAGAGACCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2461 GTCCACAGAGACCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2521 TGTGACTTGGACTGGCGGAGACATCTGAGCGACTCCAGCTAGCTGCTCAGGGGCAAC 2580
2521 TGTGACTTGGACTGGCGGAGACATCTGAGCGACTCCAGCTAGCTGCTCAGGGGCAAC 2580
2581 GCACGGCTGCGGCTGAAGTGAAGTGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
2581 GCACGGCTGCGGCTGAAGTGAAGTGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
2641 AAGAGTACGCTGCTGCTTACGGCATCTCTCTGGGATATTTTCTGAGTGTGAGATC 2700
2641 AAGAGTACGCTGCTGCTTACGGCATCTCTCTGGGATATTTTCTGAGTGTGAGATC 2700
2701 CTTTACCTTGGATCTGCTGACAGCTTATCTATTAAGTGAAGTGAAGTGAAGTGAAG 2760
2701 CTTTACCTTGGATCTGCTGACAGCTTATCTATTAAGTGAAGTGAAGTGAAGTGAAG 2760
2761 ATGGAGAGAGCTTCTATGAGAGAGAGAGAGATATCTTTGTAATGCAATCTCTGGGCT 2820
2761 ATGGAGAGAGCTTCTATGAGAGAGAGAGATATCTTTGTAATGCAATCTCTGGGCT 2820
2821 TTGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
2821 TTGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
2881 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
2881 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
2941 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
2941 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
3001 AGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3001 AGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3061 AGATTAGCTTCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
3061 AGATTAGCTTCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
3121 AGATGCTGTCTGCCATTACTCAAGATGACTTCTATAAATCAAACTCTCTCCGACAG 3180

Db 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCCGACAG 3180
QY 3181 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Db 3181 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
QY 3241 AGCTTCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 AGCTTCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
QY 3301 ACCAAACCCCTTTTCTGCTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Db 3301 ACCAAACCCCTTTTCTGCTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
QY 3361 AATACTATGTAACCTTTTCTCATCTATTTAGTGAATATTTTATGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAACCTTTTCTCATCTATTTAGTGAATATTTTATGATGGAATAAATCTTC 3420
QY 3421 TACTGTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3453
Db 3421 TACTGTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3453

RESULT 2
US-07-977-451-1
; Sequence 1, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/679,666
;; FILING DATE: 02-APR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Feit, Irving N.
;; REGISTRATION NUMBER: 28,601
;; REFERENCE/DOCKET NUMBER: LEM-3-7P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-645-1405
;; TELEFAX: 212-645-2054
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3453 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 112..3006
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 31..111
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 31..3009
US-07-977-451-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGGCTGTGCTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGAGCGACGCGCGG 60
DB 1 GCGGCTGTGCTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGAGCGACGCGCGG 60
QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 CCATCATCTGCTACCGGATGCTGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
DB 181 CCATCATCTGCTACCGGATGCTGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
DB 241 CAGAGTGAAGGACGGTATATGAAGCGGCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGTCTGCCACCCCGAGGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360
DB 301 ACCCTGCAAGTGCAGTCTGCCACCCCGAGGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420
DB 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420
QY 421 ATCTTGAACGTGTACAGAGCCCGAGGAGAACTACTTCCATATTCACAGCGCAACGC 480
DB 421 ATCTTGAACGTGTACAGAGCCCGAGGAGAACTACTTCCATATTCACAGCGCAACGC 480
QY 481 GCCAATACACGTACTGTTACAGTGAATGTGAAGATACACAGTGTATGCTGAAG 540
DB 481 GCCAATACACGTACTGTTACAGTGAATGTGAAGATACACAGTGTATGCTGAAG 540
QY 541 AGACCTTACTTTAGGAAGATGGAACCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 AGACCTTACTTTAGGAAGATGGAACCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

DB 541 AGACCTTACTTTAGGAAGATGGAACCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 CCGGAGCCCACTGTGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 CCGGAGCCCACTGTGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACAGGCTCCTCAGAGCACACTGCCCCAGTATTCTGAAAGTGGGGAAACCTTCTG 840
DB 781 AACAGGCTCCTCAGAGCACACTGCCCCAGTATTCTGAAAGTGGGGAAACCTTCTG 840
QY 841 ATCAGCTGTAAAGCCATCCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
DB 841 ATCAGCTGTAAAGCCATCCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
QY 901 AAAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTTACCTCCACAAACAGGACCATG 960
DB 901 AAAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTTACCTCCACAAACAGGACCATG 960
QY 961 ATTCGGATTCTCTTGGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 ATTCGGATTCTCTTGGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 TGCTCTTCTCAAGCACCCAGCCAGTACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TGCTCTTCTCAAGCACCCAGCCAGTACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 ATAAACGCTACCAAGTATGAAATTTGACCCGTCACGAAAGTCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 ATAAACGCTACCAAGTATGAAATTTGACCCGTCACGAAAGTCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 TTTCTCTGTGAACAGAGAGGCTGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
DB 1201 TTTCTCTGTGAACAGAGAGGCTGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
QY 1261 AAGAACAGCCAGGAGAGTACATATTTATGCAAGAAATGATGAGCCCGAGTTCACCAAA 1320
DB 1261 AAGAACAGCCAGGAGAGTACATATTTATGCAAGAAATGATGAGCCCGAGTTCACCAAA 1320
QY 1321 ATGTTTCAAGCTGAATATGAAGAAAGACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
DB 1321 ATGTTTCAAGCTGAATATGAAGAAAGACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
QY 1381 GCCTCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGGAAGAGTGTTCGGAC 1440
DB 1381 GCCTCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGGAAGAGTGTTCGGAC 1440
QY 1441 AAATCTCCCAATTTGACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500
DB 1441 AAATCTCCCAATTTGACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500
QY 1501 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTTAAATGAGTGGAGGCGGGAAGGG 1560
DB 1501 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTTAAATGAGTGGAGGCGGGAAGGG 1560
QY 1561 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1561 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 AACTCAGCGGCGCTTCCCTTTTATCCAAAGACAACTCTCTTCTATGCGACCATTTGGG 1680
DB 1621 AACTCAGCGGCGCTTCCCTTTTATCCAAAGACAACTCTCTTCTATGCGACCATTTGGG 1680

Wed May 28 09:14:08 2003

```

> CLASSIFICATION: 536
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US/07/813,593
> FILING DATE: 24-DEC-1991
> PRIOR APPLICATION DATA: US/07/793,065
> FILING DATE: 15-NOV-1991
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US/07/728,913
> FILING DATE: 28-JUN-1991
> PRIOR APPLICATION DATA: US/07/679,666
> FILING DATE: 02-APR-1991
> ATTORNEY/AGENT INFORMATION:
> NAME: Feit, Irving N.
> REGISTRATION NUMBER: 28,601
> REFERENCE/DOCKET NUMBER: LEM-3-PPP
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 212-645-1405
> TELEFAX: 212-645-2054
> INFORMATION FOR SEQ ID NO: 1:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 3453 base pairs
> TYPE: NUCLEIC ACID
> STRANDEDNESS: single
> TOPOLOGY: linear
> MOLECULE TYPE: cDNA
> FEATURE:
> NAME/KEY: CDS
> LOCATION: 31..3009
> FEATURE:
> NAME/KEY: mat_peptide
> LOCATION: 31..3006
> US-07-946-507-1

```

[illegible]

Db	481	GCCAACTACACAGTACTGTTCACAGTGAAATGTAAGAGATACACAGCTGTATGTGCTAAGG	540
Qy	541	AGACCTTACTTTAGGAAGATGAAACACAGAGTGCACCTGCTCTGCATCTCCGAGGTTGTT	600
Db	541		
Db	541	AGACCTTACTTTAGGAAGATGAAACACAGAGTGCACCTGCTCTGCATCTCCGAGGTTGTT	600
Qy	601	CGGAGCCCACTGTGGAGTGGGTGCTGCAGCTCCACAGAGAAAGCTGTAAAGAA	660
Db	601		
Db	601	CGGAGCCCACTGTGGAGTGGGTGCTGCAGCTCCACAGAGAAAGCTGTAAAGAA	660
Qy	661	GGCCCTGCTGTCTCAGAAAGCAGGAAAGAGTACTTTCATGAGTGTTCGGAACACACATC	720
Db	661		
Db	661	GGCCCTGCTGTCTCAGAAAGCAGGAAAGAGTACTTTCATGAGTGTTCGGAACACACATC	720
Qy	721	AGATGCTGTCTAGAATGCATGGGCCCGGAATGCACCAAGCTGTTTACCATAGATCTA	780
Db	721		
Db	721	AGATGCTGTCTAGAATGCATGGGCCCGGAATGCACCAAGCTGTTTACCATAGATCTA	780
Qy	781	AACACGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGTGG	840
Db	781		
Db	781	AACACGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGTGG	840
Qy	841	ATCAGGTGTAAGSCCATCCATGTGAACCATGGATTGGGCTCACCTGGGAGCTGGAAGAC	900
Db	841		
Db	841	ATCAGGTGTAAGSCCATCCATGTGAACCATGGATTGGGCTCACCTGGGAGCTGGAAGAC	900
Qy	901	AAAGCCCTGGAGAGGCGAGCTACTTTGAGTAGTAGTACTTCCACAAACAGGACCATG	960
Db	901		
Db	901	AAAGCCCTGGAGAGGCGAGCTACTTTGAGTAGTAGTACTTCCACAAACAGGACCATG	960
Qy	961	ATTCCGATTCCTTTGGCCTTTGTGCTTCCGTGGGGAAGAACGACACCCGGATATACACC	1020
Db	961		
Db	961	ATTCCGATTCCTTTGGCCTTTGTGCTTCCGTGGGGAAGAACGACACCCGGATATACACC	1020
Qy	1021	TGCTCTTCCTCAAGCACCCACGACAGTCAGCGTTGCTGACCATCTCAGAAAAGGTTT	1080
Db	1021		
Db	1021	TGCTCTTCCTCAAGCACCCACGACAGTCAGCGTTGCTGACCATCTCAGAAAAGGTTT	1080
Qy	1081	ATAAAGCTACCAAGCTCGCAAGAGAGTAGTAAATTTGACCCGTACGAAAGTTCTGCTTC	1140
Db	1081		
Db	1081	ATAAAGCTACCAAGCTCGCAAGAGAGTAGTAAATTTGACCCGTACGAAAGTTCTGCTTC	1140
Qy	1141	TCAGTCAGGTTTAAAGGTACCCACAGAAATCCGATGCACGTGGATCTTCTCAAGCCTCA	1200
Db	1141		
Db	1141	TCAGTCAGGTTTAAAGCGTACCCACAGAAATCCGATGCACGTGGATCTTCTCAAGCCTCA	1200
Qy	1201	TTTCCTTGTCAACAGAGAGCCGTGGAGTGGGTACAGCATATCTAAATTTTTCGATCAT	1260
Db	1201		
Db	1201	TTTCCTTGTGAACAGAGAGCCGTGGAGTGGGTACAGCATATCTAAATTTTTCGATCAT	1260
Qy	1261	AAGAACAAGCCAGGAGAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTTCAACAA	1320
Db	1261		
Db	1261	AAGAACAAGCCAGGAGAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTTCAACAA	1320
Qy	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCCTACCAATGCCCTCAGCCAG	1380
Db	1321		
Db	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCCTACCAATGCCCTCAGCCAG	1380
Qy	1381	CGCTCTGTTCTCTGATGCTACCCGCTACCCCTTTGGACCTGTGAAGAGTGTTCGGAC	1440
Db	1381		
Db	1381	CGCTCTGTTCTCTGATGCTACCCGCTACCCCTTTGGACCTGTGAAGAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA	1500
Db	1441		
Db	1441	AAATCTCCCAATTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA	1500
Qy	1501	AAAGTGTTTGGCCAGTGGGTGTCGACAGTACTCTAAATATGATGAGCCCGGGAAGGG	1560
Db	1501		
Db	1501	AAAGTGTTTGGCCAGTGGGTGTCGACAGTACTCTAAATATGATGAGCCCGGGAAGGG	1560
Qy	1561	CTTCTGGTCAAAATGCTGTGGCTACAAATCTATGGGCACGCTTTGGCAAAACATCTTTTA	1620
Db	1561		
Db	1561	CTTCTGGTCAAAATGCTGTGGCTACAAATCTATGGGCACGCTTTGGCAAAATCTTTTA	1620

```
QY 1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGACAACATCTCTCTATCGACCATTTGG 1680
DB 1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGACAACATCTCTCTATCGACCATTTGG 1680
QY 1681 CTCCTCTCCCTTCCATGTTGTTCTCATTTGTTGATGTCGCCAATAACAAAACCAA 1740
DB 1681 CTCCTCTCCCTTCCATGTTGTTCTCATTTGTTGATGTCGCCAATAACAAAACCAA 1740
QY 1741 TTTAGGTACGAGAGCTCAGCTGAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTAC 1800
DB 1741 TTTAGGTACGAGAGCTCAGCTGAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTAC 1800
QY 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860
DB 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860
QY 1861 TTAGAGTTTGGGAAGGCTCTGGGGTCTGGCGCTTTCCGGGAGGGTGATGAACCCAGGCC 1920
DB 1861 TTAGAGTTTGGGAAGGCTCTGGGGTCTGGCGCTTTCCGGGAGGGTGATGAACCCAGGCC 1920
QY 1921 TATGSCATTAGTAAACGGAGTCTCAATTGAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
DB 1921 TATGSCATTAGTAAACGGAGTCTCAATTGAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTGAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
DB 1981 GCTGACAGCTGTGAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100
DB 2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100
QY 2101 ATTTTGAATATTGTTGATGGTGACCTCCTCACTACCTAAGAGTAAAGAGAGAG 2160
DB 2101 ATTTTGAATATTGTTGATGGTGACCTCCTCACTAAGAGTAAAGAGAGAG 2160
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATATTTTCAGTTCTTACCTACTTTC 2220
DB 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATATTTTCAGTTCTTACCTACTTTC 2220
QY 2221 CAGGCACATTCAAATTCAGCATGCTGGTTCCAGGAGAGTTTCAGTTACACCCGCCCTTG 2280
DB 2221 CAGGCACATTCAAATTCAGCATGCTGGTTCCAGGAGAGTTTCAGTTACACCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATTCGAGATGAGATGAATGAA 2340
DB 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATTCGAGATGAGATGAATGAA 2340
QY 2341 AACCAGAGAGGCTGCGAGAGAGAGAGAGATTTGAACGTCGTCGAGTTTGAAGAC 2400
DB 2341 AACCAGAGAGGCTGCGAGAGAGAGAGAGATTTGAACGTCGTCGAGTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTCGCTACCACTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCGCT 2460
DB 2401 CTCCTTTGCTTTCGCTACCACTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCGCT 2460
QY 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTACCCCGGAAGGTGTTGAAGATC 2520
DB 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTACCCCGGAAGGTGTTGAAGATC 2520
QY 2521 TGTGACTTGGACTGGCCCGAGACATCCTGAGGACTCCAGCTACGTCGTCAGGGGCAC 2580
DB 2521 TGTGACTTGGACTGGCCCGAGACATCCTGAGGACTCCAGCTACGTCGTCAGGGGCAC 2580
QY 2581 GCACGGCTGCCGTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACAATC 2640
DB 2581 GCACGGCTGCCGTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACAATC 2640
QY 2641 AAGAGTGAGCTGTGCTCTACGGCATCTCTCTCTGGAGATATTTTCACTGGGTGTGAC 2700
DB 2641 AAGAGTGAGCTGTGCTCTACGGCATCTCTCTCTGGAGATATTTTCACTGGGTGTGAC 2700
```

```
QY 2701 CCTTACCCTGGCATTCCTGTGACGCTAACTTCTATAAACTGATTGAGTGGATTTAAA 2760
DB 2701 CCTTACCCTGGCATTCCTGTGACGCTAACTTCTATAAACTGATTGAGTGGATTTAAA 2760
QY 2761 ATGGACAGCCATTCATGCCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
DB 2761 ATGGACAGCCATTCATGCCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACTGACTTCTATTTTAGGATGTCAGCTG 2880
DB 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACTGACTTCTATTTTAGGATGTCAGCTG 2880
QY 2881 GCAGAGCGCAGAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGCGGCCCT 2940
DB 2881 GCAGAGCGCAGAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGCGGCCCT 2940
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCCAGTCGCCACAGCGCCAGGTCGAGATTCACAGACAA 3000
DB 2941 CAGCAGAGAGCGGGCTCAGAGCCCAGTCGCCACAGCGCCAGGTCGAGATTCACAGACAA 3000
QY 3001 AGAAGTTAGCGAGCAGCGCTTGGACCCCGCCCTACACGCTGTAGCCGCGACGCCA 3060
DB 3001 AGAAGTTAGCGAGCAGCGCTTGGACCCCGCCCTACACGCTGTAGCCGCGACGCCA 3060
QY 3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGCTTCTGCTGCTGGACTTTCTCT 3120
DB 3061 AGATTAGCTCGCTCTGAGGAAGCGCTTACAGCGCTTCTGCTGCTGGACTTTCTCT 3120
QY 3121 AGATGCTGTGCTGCTTACTCCTCAAGTGACTTCTATAAAATCAAACCTCTCTCGCACAG 3180
DB 3121 AGATGCTGTGCTGCTTACTCCTCAAGTGACTTCTATAAAATCAAACCTCTCTCGCACAG 3180
QY 3181 GCGGAGAGCCATATAGACTTGTGGTGACCGCCCTACCTGCGGGCCCTTTCACG 3240
DB 3181 GCGGAGAGCCATATAGACTTGTGGTGACCGCCCTACCTGCGGGCCCTTTCACG 3240
QY 3241 AGCTTCAGGGGAAGCATGTATCTGAAATATAGTATATTTCTGTAATAGTGAACAA 3300
DB 3241 AGCTTCAGGGGAAGCATGTATCTGAAATATAGTATATTTCTGTAATAGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTGTCTAAGGGAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360
DB 3301 ACCAAACCCGTTTTTGTCTAAGGGAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360
QY 3361 AATACTATGTAATCTTTTCTATCTATTTAGTGATATATTTTATGATGGAATAAATTC 3420
DB 3361 AATACTATGTAATCTTTTCTATCTATTTAGTGATATATTTTATGATGGAATAAATTC 3420
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453
DB 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453
```

RESULT 4

```
US-08-252-517-1
; Sequence 1, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


Db 1201 TTCTCTGTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTTCGGATCAT 1260
QY 1261 AAGAAACAGCCAGGAGGTACATATCTATGAGAAAATGATGACCCCGAGTTCCACCAA 1320
Db 1261 AAGAAACAGCCAGGAGGTACATATCTATGAGAAAATGATGACCCCGAGTTCCACCAA 1320
QY 1321 ATGTTACAGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATCCCTCAGCCAGCCAG 1380
Db 1321 ATGTTACAGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATCCCTCAGCCAGCCAG 1380
QY 1381 GCGTCTGTCTCTGTATGGCTACCGCTACCGCTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440
Db 1381 GCGTCTGTCTCTGTATGGCTACCGCTACCGCTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440
QY 1441 AATCTCCCAATGTCAGGAGGAAATCCAGAGGAGTTTGAATAAAAAGGCTAACAGA 1500
Db 1441 AATCTCCCAATGTCAGGAGGAAATCCAGAGGAGTTTGAATAAAAAGGCTAACAGA 1500
QY 1501 AAGTGTGTTGGCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAGGCGCGGAAGGG 1560
Db 1501 AAGTGTGTTGGCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAGGCGCGGAAGGG 1560
QY 1561 CTTCTGGTCAATGCTGCGTACAAATCTATGGCACGCTTTGGAAACCAATCTTTTAA 1620
Db 1561 CTTCTGGTCAATGCTGCGTACAAATCTATGGCACGCTTTGGAAACCAATCTTTTAA 1620
QY 1621 AACTCAGAGGCCCTTCCCTTTCATCCAGACACATCTCTATCGGACCATGGG 1680
Db 1621 AACTCAGAGGCCCTTCCCTTTCATCCAGACACATCTCTATCGGACCATGGG 1680
QY 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
QY 1741 TTTAGGTAGAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCTCGATACAGAGTAC 1800
Db 1741 TTTAGGTAGAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCTCGATACAGAGTAC 1800
QY 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCGAGAGAGAC 1860
Db 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCGAGAGAGAC 1860
QY 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTTCGGAGGGTGATGAACGCCAGCGGC 1920
Db 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTTCGGAGGGTGATGAACGCCAGCGGC 1920
QY 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTAAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTGTAAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100
QY 2101 ATTTTGAATATGTTGCTATGGTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
Db 2101 ATTTTGAATATGTTGCTATGGTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160
QY 2161 TTTACAGGACATGACAGAGATTTTAAAGAACATATTTTACGTTCTTACCCCTACTTTC 2220
Db 2161 TTTACAGGACATGACAGAGATTTTAAAGAACATATTTTACGTTCTTACCCCTACTTTC 2220
QY 2221 CAGGCACATTCAAATTCAGCATGCTGCTTCCAGGAGAGTTTCAAGTACACCCCGCCCTTG 2280
Db 2221 CAGGCACATTCAAATTCAGCATGCTGCTTCCAGGAGAGTTTCAAGTACACCCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGGTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2340

QY 2341 AACAGAAGAGCTGGCAGAGAAGAGGAGGAAGATTTGAACGTGCTGACGTTTGAAGAC 2400
Db 2341 AACAGAAGAGCTGGCAGAGAAGAGGAGGAAGATTTGAACGTGCTGACGTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTCGTCACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTG 2460
Db 2401 CTCCTTTGCTTTCGTCACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTG 2460
QY 2461 GTCCACAGAGACCTGGCAGGCCAGGAATGTTGGTCAACCCACGCCGGAAGGTTGGAATC 2520
Db 2461 GTCCACAGAGACCTGGCAGGCCAGGAATGTTGGTCAACCCACGCCGGAAGGTTGGAATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGGAGACATCCTCGAGGACATCCAGTACGTCGTGAGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGGAGACATCCTCGAGGACATCCAGTACGTCGTGAGGGCAAC 2580
QY 2581 GCACGGCTGCCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGATCTACACAATC 2640
Db 2581 GCACGGCTGCCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGATCTACACAATC 2640
QY 2641 AAGAGTGACGTCTGTGCTCTACGGCATCCTCTCTGGGAGATATTTTTCACGTTGTGAAC 2700
Db 2641 AAGAGTGACGTCTGTGCTCTACGGCATCCTCTCTGGGAGATATTTTTCACGTTGTGAAC 2700
QY 2701 CCTTACCTGGCATTCCTGTGCGAGCTAACTTCTATAACTGATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTGGCATTCCTGTGCGAGCTAACTTCTATAACTGATTCAGAGTGGATTTAAA 2760
QY 2761 ATGAGCAGCATTCTATGCCACAGAGGATATACCTTGTGTAATGCAATCCTGCTGGGCT 2820
Db 2761 ATGAGCAGCATTCTATGCCACAGAGGATATACCTTGTGTAATGCAATCCTGCTGGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCCATCCTCCCAAGCTGACTTCTTTTAGGATGTGAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCCATCCTCCCAAGCTGACTTCTTTTAGGATGTGAGCTG 2880
QY 2881 GCAGAGCAGAAAGACATGATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGCAGAAAGACATGATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940
QY 2941 CACAGAGAGCGGGCTCAGAGCCAGTCCGACAGCCAGCCAGGTGAAGATTCACAGAGAA 3000
Db 2941 CACAGAGAGCGGGCTCAGAGCCAGTCCGACAGCCAGCCAGGTGAAGATTCACAGAGAA 3000
QY 3001 AGAAGTTAGCGAGAGCGCTTTGAGCCGCCACCCCTAGCAGGTGTAGACGCGAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGAGCGCTTTGAGCCGCCACCCCTAGCAGGTGTAGACGCGAGAGCCA 3060
QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120
Db 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120
QY 3121 AGATGCTGTGCTACTTCTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCAGAG 3180
Db 3121 AGATGCTGTGCTACTTCTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCAGAG 3180
QY 3181 GCGGGAGAGCAATAATAGACTTGTGTGAGCCGCCCTACCCCTGGGGGCTTTTTCACG 3240
Db 3181 GCGGGAGAGCAATAATAGACTTGTGTGAGCCGCCCTACCCCTGGGGGCTTTTTCACG 3240
QY 3241 ACCTTACGGGAAGCCATGATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300
Db 3241 ACCTTACGGGAAGCCATGATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300
QY 3301 ACCAAACCGTTTTTTCCTAAGGGAAGCTTAAATATGATTTTAAAAATCTATCTTTTAA 3360
Db 3301 ACCAAACCGTTTTTTCCTAAGGGAAGCTTAAATATGATTTTAAAAATCTATCTTTTAA 3360
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGGATGGAATAAATCTTC 3420

OY 3421 TACTGTAAACAAAAA 3453
Db 3421 TACTGTAAACAAAAA 3453
RESULT 5
US-07-906-397A-1
; Sequence 1, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..3009
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 31..3006
; US-07-906-397A-1
Query Match 100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCGGCTTGGCTACCGCGCTCCGGAGGCGCATCGGGCGTTCGGCGAGCGCAGCGACCGG 60
|||||
61 CGGCTGCTGCTGCTTGTGTTTGTTCAGTAATGATTCTTGAGACCGTTACAAACCAAGAC 120
|||||
61 CGGCTGCTGCTGCTTGTGTTTGTTCAGTAATGATTCTTGAGACCGTTACAAACCAAGAC 120
|||||
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAATGGCTCATCAGCGGGAAG 180
|||||
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAATGGCTCATCAGCGGGAAG 180
|||||
181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTACCCCGAGCGCG 240
|||||
181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTACCCCGAGCGCG 240
|||||
241 CAGAGTGAAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
|||||
241 CAGAGTGAAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
|||||
301 ACCCTGCAAGTGCAGTCCGCCACCGGAGGACCTTCTCTGCTCTGGGTCTTTAAGCAC 360
|||||
301 ACCCTGCAAGTGCAGTCCGCCACCGGAGGACCTTCTCTGCTCTGGGTCTTTAAGCAC 360
|||||
361 AGCTCCCTGGGCTGCGAGCGGCACCTTTGATTTACAAAACAGAGAGAAATCGTTTCCATGGCC 420
|||||
361 AGCTCCCTGGGCTGCGAGCGGCACCTTTGATTTACAAAACAGAGAGAAATCGTTTCCATGGCC 420
|||||
421 ATCTTGAACGTGACAGAGACCGAGGAGGAGTAATCTTCTCTCATATTCAGAGCGAACGC 480
|||||
421 ATCTTGAACGTGACAGAGACCGAGGAGGAGTAATCTTCTCTCATATTCAGAGCGAACGC 480
|||||
481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
|||||
481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
|||||
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTGCTGCTCTCCGAGGCTGTT 600
|||||
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTGCTGCTCTCCGAGGCTGTT 600
|||||
601 CCGGAGCCCTGCTGAGTGGGTGCTCTGAGTCTCCACAGGAGAAAGCTGTAAAGAGAA 660
|||||
601 CCGGAGCCCTGCTGAGTGGGTGCTCTGAGTCTCCACAGGAGAAAGCTGTAAAGAGAA 660
|||||
661 GGCCCTGCTGTTGTCAGAAAGAGAGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720
|||||
661 GGCCCTGCTGTTGTCAGAAAGAGAGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720
|||||
721 AGATGCTGTCTAGAAATGCAGTGGCGCGAATGCAACAGCTGTTCACCATAGATCTA 780
|||||
721 AGATGCTGTCTAGAAATGCAGTGGCGCGAATGCAACAGCTGTTCACCATAGATCTA 780
|||||
781 AACAGGCTCCTCAGAGCACACTGCCCCAGTTATCTCTGAAAGTGGGGGAAACCCCTTGTGG 840
|||||
781 AACAGGCTCCTCAGAGCACACTGCCCCAGTTATCTCTGAAAGTGGGGGAAACCCCTTGTGG 840
|||||
841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900
|||||
841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900
|||||
901 AAAGCCCTGGAGGAGGCGACGCTACTTTGAGATGAGTACCTACTCCACAAACAGAGACCATG 960
|||||
901 AAAGCCCTGGAGGAGGCGACGCTACTTTGAGATGAGTACCTACTCCACAAACAGAGACCATG 960
|||||
961 ATTCGGATTCCTTGGCCCTTGTGCTCTCCGTTGGAGGAGGAGGAGGATATACACC 1020
|||||
961 ATTCGGATTCCTTGGCCCTTGTGCTCTCCGTTGGAGGAGGAGGAGGATATACACC 1020
|||||
1021 TGCTCTTCTCTAAAGCACCCAGCAGCTAGCGTGTGGTGAACCTCTCTAGAAAAAGGGTTT 1080
|||||
1021 TGCTCTTCTCTAAAGCACCCAGCAGCTAGCGTGTGGTGAACCTCTCTAGAAAAAGGGTTT 1080
|||||
1081 ATAAAGCTACAGCTCGAAGAGAGATGAAATTGACCCGTTACGAAAGTCTTCGCTTC 1140
|||||
1081 ATAAAGCTACAGCTCGAAGAGAGATGAAATTGACCCGTTACGAAAGTCTTCGCTTC 1140
|||||

OY 1 GCGGCTTGGCTACCGCGCTCCGGAGGCGCATCGGGCGTTCGGCGAGCGCAGCGACCGG 60
|||||

1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA 1200
1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA 1200
1201 TTTCTCTGTGAACAGAGAGGCGCTGGAGGTGATACAGATATCTAAATTTTGGCATCAT 1260
1201 TTTCTCTGTGAACAGAGAGGCGCTGGAGGTGATACAGATATCTAAATTTTGGCATCAT 1260
1261 AAGAACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCAGTTTCAACCAA 1320
1261 AAGAACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCAGTTTCAACCAA 1320
1321 ATGTTACAGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCCTCAGCCAGCCAG 1380
1321 ATGTTACAGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCCTCAGCCAGCCAG 1380
1381 GCGTCTCTCTGATGGCTACCGCTACCCCTTACCTCTTGGACCTGGAAAGGTGTTGGGAC 1440
1381 GCGTCTCTCTGATGGCTACCGCTACCCCTTACCTCTTGGACCTGGAAAGGTGTTGGGAC 1440
1441 AAATCTCCCAATTGCACGAGGAATCCAGAGAGGTTTGGAAATAAAGGCTACACAGA 1500
1441 AAATCTCCCAATTGCACGAGGAATCCAGAGAGGTTTGGAAATAAAGGCTACACAGA 1500
1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGGCGGGAAAGGG 1560
1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGTGAGGCGGGAAAGGG 1560
1561 CTCTCGGTCAATGCTGTGCGTACAAATCTATGGGACGCTTGGGAAACCATCTTTTAA 1620
1561 CTCTCGGTCAATGCTGTGCGTACAAATCTATGGGACGCTTGGGAAACCATCTTTTAA 1620
1621 AACTCAGAGGCCCCCTTCCCTTTCATGCAAGACACATCTCTTATGCGGACCATTTGGG 1680
1621 AACTCAGAGGCCCCCTTCCCTTTCATGCAAGACACATCTCTTATGCGGACCATTTGGG 1680
1681 CTCTCTCTCCCTTCTATGTTCTCTCATTTGTTGATCTGCCAATAACAAAAGCAA 1740
1681 CTCTCTCTCCCTTCTATGTTCTCTCATTTGTTGATCTGCCAATAACAAAAGCAA 1740
1741 TTTAGTAGGAGAGTACAGTGCAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTAC 1800
1741 TTTAGTAGGAGAGTACAGTGCAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTAC 1800
1801 TTCTAGTTGACTTCAGGAGCTATGAATATGACCTTTAAGTGGAGTTTCCCGAGAGAGAAC 1860
1801 TTCTAGTTGACTTCAGGAGCTATGAATATGACCTTTAAGTGGAGTTTCCCGAGAGAGAAC 1860
1861 TTAGAGTTTGGGAAGTCTGGGCTGGCGCTTTTCGGGAGGTTGATGAACGCCAGCGCC 1920
1861 TTAGAGTTTGGGAAGTCTGGGCTGGCGCTTTTCGGGAGGTTGATGAACGCCAGCGCC 1920
1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
1981 GCTGACAGCTGTGAAAAGAACTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGCTGTGAAAAGAACTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACAACTGTAATCTGTCGGGGGATGCACACTGTGACGGGCGAGTGTACTTGG 2100
2041 CACCATGACAACTGTAATCTGTCGGGGGATGCACACTGTGACGGGCGAGTGTACTTGG 2100
2101 ATTTTGTGAATATGTTGCTATGTTGACCTCCTCAACTACTTAAGAAATGAAAGAGAGAAG 2160
2101 ATTTTGTGAATATGTTGCTATGTTGACCTCCTCAACTACTTAAGAAATGAAAGAGAGAAG 2160
2161 TTTTCAGGACATGGACAGAGATTTTAAAGAAACATAATTTTTCAGTTCTTACCCCTACTTTC 2220
2161 TTTTCAGGACATGGACAGAGATTTTAAAGAAACATAATTTTTCAGTTCTTACCCCTACTTTC 2220

2221 CAGGCACATTCAAAATCCAGCATGCTCGTGTTCACGAGAAGTTTCAGTTACACCGCCCTTG 2280
2221 CAGGCACATTCAAAATCCAGCATGCTCGTGTTCACGAGAAGTTTCAGTTACACCGCCCTTG 2280
2281 GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCCTCAAGATGAGATTGAATGAA 2340
2281 GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCCTCAAGATGAGATTGAATGAA 2340
2341 AACCAGAAAGGCTGCAGAAAGAGAGAGATTGACGTCGTCGAGCTTGGAGAC 2400
2341 AACCAGAAAGGCTGCAGAAAGAGAGAGATTGACGTCGTCGAGCTTGGAGAC 2400
2401 CTCTTGTGCTTTCGCTACCAAGTGGCAAGGATGAAATTCCTGGAGTTCAAGTCGTGT 2460
2401 CTCTTGTGCTTTCGCTACCAAGTGGCAAGGATGAAATTCCTGGAGTTCAAGTCGTGT 2460
2461 GTCCACAGAGACTGTCAGCCAGGAATGTTGTCACCCAGGAAAGGTGGTGAAGATC 2520
2461 GTCCACAGAGACTGTCAGCCAGGAATGTTGTCACCCAGGAAAGGTGGTGAAGATC 2520
2521 TGTGACTTTGGAGTGGCCGAGACATCCTCAGCGACTCCAGCTACGTCGTCAGSGGCAAC 2580
2521 TGTGACTTTGGAGTGGCCGAGACATCCTCAGCGACTCCAGCTACGTCGTCAGSGGCAAC 2580
2581 GCACGGCTGCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC 2640
2581 GCACGGCTGCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC 2640
2641 AAGAGTACGCTGTGCTTACGGCATCTTCTCTGGAGATATTTTTCACCTGGGTGTGAAC 2700
2641 AAGAGTACGCTGTGCTTACGGCATCTTCTCTGGAGATATTTTTCACCTGGGTGTGAAC 2700
2701 CCTTACCTTGGCATTCCTGTCGAGCTTAATCTTATATAACTGATTCAGAGTGGATTTAAA 2760
2701 CCTTACCTTGGCATTCCTGTCGAGCTTAATCTTATATAACTGATTCAGAGTGGATTTAAA 2760
2761 ATGGAGCAGCTATCTATGCCACAGAGGATATACTTTGTAATGCAATCCTGCTGGCT 2820
2761 ATGGAGCAGCTATCTATGCCACAGAGGATATACTTTGTAATGCAATCCTGCTGGCT 2820
2821 TTTGACTCAAGGAAGCGGCTATCCTTCCCACTGACTTCTATTTTAGGATGTCAGCTG 2880
2821 TTTGACTCAAGGAAGCGGCTATCCTTCCCACTGACTTCTATTTTAGGATGTCAGCTG 2880
2881 GCAGAGCAGAAAGATGATCAGAAATCCTATCCTACCAAGAGGCGGCCCT 2940
2881 GCAGAGCAGAAAGATGATCAGAAATCCTATCCTACCAAGAGGCGGCCCT 2940
2941 CAGCAGAGGCGGCTCAGAGCCAGTCCAGCGCCAGTGAAGATTTCACAGAGAA 3000
2941 CAGCAGAGGCGGCTCAGAGCCAGTCCAGCGCCAGTGAAGATTTCACAGAGAA 3000
3001 AGAAGTTAGCGAGGCGCTTGGACCCGCTAGCAGCTAGCAGCTGAGAGCCGAGAGCCA 3060
3001 AGAAGTTAGCGAGGCGCTTGGACCCGCTAGCAGCTAGCAGCTGAGAGCCGAGAGCCA 3060
3061 AGATTAGCTCGCTCTGAGGAAGCGGCTACAGCGGTTGCTTGGAGCTTCTCT 3120
3061 AGATTAGCTCGCTCTGAGGAAGCGGCTACAGCGGTTGCTTGGAGCTTCTCTCT 3120
3121 AGATGCTGTGCTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180
3121 AGATGCTGTGCTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180
3181 GCGGAGAGCCAAATATGAGACTTGTGTTGAGCCGCTACCTTGGGCGCTTTTCCACG 3240
3181 GCGGAGAGCCAAATATGAGACTTGTGTTGAGCCGCTACCTTGGGCGCTTTTCCACG 3240
3241 AGCTTGAGGGAAGCCATGATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
3241 AGCTTGAGGGAAGCCATGATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
3301 ACCAAACCCGTTTTTGTCTAAGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA 3360

Db	3301	ACCAACCCGTTTTTTCCTAAGGGAAGCTAAATATGATTTTTAAAAATCTATGTTTAA	3360
Qy	3361	AATACATGTAACCTTTTTCATCTATTAGTGATATATTTATGATGGAATAAACTTTC	3420
Db	3361	AATACATGTAACCTTTTTCATCTATTAGTGATATATTTATGATGGAATAAACTTTC	3420
Qy	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453
Db	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

```

RESULT 6
US-08-601-891-1
; Sequence 1, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,891
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; *PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGACCAAGCTGTTCACCATATCTA 780
QY 781 AACAGGCTCTCAGAGCACACTGCCAGATTATTCCTGAAGTGGGGAACCCCTGTGG 840
Db 781 AACAGGCTCTCAGAGCACACTGCCAGATTATTCCTGAAGTGGGGAACCCCTGTGG 840
QY 841 ATCAGGTGTAGGCCATCATGTGAACCATGTGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAGGCCATCATGTGAACCATGTGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
QY 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTACCAACAGGACCATG 960
Db 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTACCAACAGGACCATG 960
QY 961 ATTCGGATTCTCTGGCCCTTTCTGCTCCGTTGGGAAGAACGACACCGGATATTACAC 1020
Db 961 ATTCGGATTCTCTGGCCCTTTCTGCTCCGTTGGGAAGAACGACACCGGATATTACAC 1020
QY 1021 TGCTCTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAAGGTTT 1080
Db 1021 TGCTCTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAAGGTTT 1080
QY 1081 ATAAAGCTTACAGCTCGCAAGAGATGAAATGACCCGTACGCAAAAAGTCTGCTTC 1140
Db 1081 ATAAAGCTTACAGCTCGCAAGAGATGAAATGACCCGTACGCAAAAAGTCTGCTTC 1140
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCACGTGATCTCTCAAGCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCACGTGATCTCTCAAGCCTCA 1200
QY 1201 TTTCTCTTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
Db 1201 TTTCTCTTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
QY 1261 AAGAAACAGCCAGGAGATACATATTTCTATGCAAGAAATGATGACCCGATTCACCAA 1320
Db 1261 AAGAAACAGCCAGGAGATACATATTTCTATGCAAGAAATGATGACCCGATTCACCAA 1320
QY 1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCTCAGCCAGCCAG 1380
QY 1381 GCGTCTCTCTGATGGCTACCGCTACCCCTCTTGGACCTTGGAGAGTGTTCGGAC 1440
Db 1381 GCGTCTCTCTGATGGCTACCGCTACCCCTCTTGGACCTTGGAGAGTGTTCGGAC 1440
QY 1441 AAATCTCCAAATGTCACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
Db 1441 AAATCTCCAAATGTCACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
QY 1501 AAAGTGTGTTGGCCAGTGGGTGCGACAGTACTCTAAATATGATGAGCCGCGGAAGG 1560
Db 1501 AAAGTGTGTTGGCCAGTGGGTGCGACAGTACTCTAAATATGATGAGCCGCGGAAGG 1560
QY 1561 CTTCTGCTCAATGCTGCGTACAATTTCTATGGGCACGCTTTCGGAACCAATCTTTT 1620
Db 1561 CTTCTGCTCAATGCTGCGTACAATTTCTATGGGCACGCTTTCGGAACCAATCTTTT 1620
QY 1621 AAATCAAGGCGCCCTTCCCTTTCAATCCCAAGCAACATCTCTTATGCGACCATTTGG 1680
Db 1621 AAATCAAGGCGCCCTTCCCTTTCAATCCCAAGCAACATCTCTTATGCGACCATTTGG 1680
QY 1681 CTTCTGCTCCCTTCAATTTGTTCTCAATTTGTTGATCTGATCTGCAAAATACAAAGCAA 1740
Db 1681 CTTCTGCTCCCTTCAATTTGTTCTCAATTTGTTGATCTGATCTGCAAAATACAAAGCAA 1740
QY 1741 TTTAGTACAGAGTACAGTGCAGATGATCCAGTACTGCGCCCTTGGATACAGGATAC 1800
Db 1741 TTTAGTACAGAGTACAGTGCAGATGATCCAGTACTGCGCCCTTGGATACAGGATAC 1800
QY 1801 TTTCTAGCTTACAGGAGTATGATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTTCTAGCTTACAGGAGTATGATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860

QY 1861 TTAGAGTTTGGGAAGGTCTGGGCTCTGGCGCTTTTCGGGAGGGTGTATGAACGCCACGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCTGGGCTCTGGCGCTTTTCGGGAGGGTGTATGAACGCCACGCC 1920
QY 1921 TATGGCATTTAGTAAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGTGTCTAAAGAGAAA 1980
Db 1921 TATGGCATTTAGTAAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGTGTCTAAAGAGAAA 1980
QY 1981 GCTGACAGTGTCAAAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGTGTCAAAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACAAACATCTGTAATCTGCTGGGGCATGCACACTGTCTAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACAAACATCTGTAATCTGCTGGGGCATGCACACTGTCTAGGGCCAGTGTACTTG 2100
QY 2101 ATTTTGAATATTGTTGCTATGCTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAA 2160
Db 2101 ATTTTGAATATTGTTGCTATGCTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAA 2160
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTACCCCTTTC 2220
Db 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTACCCCTTTC 2220
QY 2221 CAGGCACTTCAAAATCCAGCATGCTGTTTACGAGAAAGTTCAAGTTACACCCGCCCTTG 2280
Db 2221 CAGGCACTTCAAAATCCAGCATGCTGTTTACGAGAAAGTTCAAGTTACACCCGCCCTTG 2280
QY 2281 GATCAGCTCTCAGGGTTCAATGGAAATTCATTTCAATCTCAAGATGAGATTTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGAAATTCATTTCAATCTCAAGATGAGATTTGAATATGAA 2340
QY 2341 AACCAAGAGGCTGCGCAGAGAGAGAGAGATTTGAACGTGTCAGCTTTTGAAGAC 2400
Db 2341 AACCAAGAGGCTGCGCAGAGAGAGAGAGATTTGAACGTGTCAGCTTTTGAAGAC 2400
QY 2401 CTCTTCTGCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
Db 2401 CTCTTCTGCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
QY 2461 GTCACAGAGACCTGCGCAGAGATGTTGTTGTCACCCAGGAGGAGGTGGTGAAGATC 2520
Db 2461 GTCACAGAGACCTGCGCAGAGATGTTGTTGTCACCCAGGAGGAGGTGGTGAAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCCACTCCAGTACGTCGTCAGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCCTGAGCCACTCCAGTACGTCGTCAGGGCAAC 2580
QY 2581 GCACGGCTGCGGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Db 2581 GCACGGCTGCGGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
QY 2641 AAGAGTGAACCTCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGAAC 2700
Db 2641 AAGAGTGAACCTCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGAAC 2700
QY 2701 CCTTACCTGGCATCTGCTGAGGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTGGCATCTGCTGAGGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760
QY 2761 ATGAGCAGCATTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820
Db 2761 ATGAGCAGCATTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGCCATCTTCCCAACCTGACTTCTATTTTATAGTGTGAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGCCATCTTCCCAACCTGACTTCTATTTTATAGTGTGAGCTG 2880
QY 2881 GCAGAGCAGAGAGAGATGATCAGAACATCTCATCTACCAAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGCAGAGAGAGATGATCAGAACATCTCATCTACCAAAAACAGGCGGCCCT 2940

|||||
361 AGCTCCCTGGGTGCGAGCGCACCTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
QY
421 ATCTTTAAAGCTCACAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACCC 480
Db
421 ATCTTTAAAGCTGACAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACCC 480
QY
481 GCCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
Db
481 GCCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
QY
541 AGACCTTACTTTAGGAAGATGAAACAGGATGACACTCTGTCATCTCCGAGGGTGT 600
Db
541 AGACCTTACTTTAGGAAGATGAAACAGGATGACACTCTGTCATCTCCGAGGGTGT 600
QY
601 CCGGAGCCCACTGTCAGTGGGTGCTCTGCAGCTCCACAGGAAGCTGTAAGAGAA 660
Db
601 CCGGAGCCCACTGTCAGTGGGTGCTCTGCAGCTCCACAGGAAGCTGTAAGAGAA 660
QY
661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTTCATGAGTGTTCGGAACACAGATC 720
Db
661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTTCATGAGTGTTCGGAACACAGATC 720
QY
721 AGATGCTGTGTAGAAATGCACTGGGCCCGGAATGCACCAAGCTGTTCAACCATAGATCTA 780
Db
721 AGATGCTGTGTAGAAATGCACTGGGCCCGGAATGCACCAAGCTGTTCAACCATAGATCTA 780
QY
781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840
Db
781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840
QY
841 ATCAGTGTAAAGCCATCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Db
841 ATCAGTGTAAAGCCATCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
QY
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTCACAAACAGGACCATG 960
Db
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTCACAAACAGGACCATG 960
QY
961 ATTCGGATCTCTGGGCCCTTTGTGCTCCGTTGGGAAGAACAGACCGGATATTACACC 1020
Db
961 ATTCGGATCTCTGGGCCCTTTGTGCTCCGTTGGGAAGAACAGACCGGATATTACACC 1020
QY
1021 TGCTCTCTCAAGACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAGGTTT 1080
Db
1021 TGCTCTCTCAAGACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAGGTTT 1080
QY
1081 ATAAACGCTACCAAGCTCGCAAGAGAGATGAAATTGACCCGTACGAAAAGTTCTGCTTC 1140
Db
1081 ATAAACGCTACCAAGCTCGCAAGAGAGATGAAATTGACCCGTACGAAAAGTTCTGCTTC 1140
QY
1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCGATGCGATCTCTCAAGCCCTCA 1200
Db
1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCGATGCGATCTCTCAAGCCCTCA 1200
QY
1201 TTTCTCTGTAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Db
1201 TTTCTCTGTAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
QY
1261 AAGAACAGCCAGGAGATACATATCTATGCAGAAAATGATGACGCCCATGTTACACAAA 1320
Db
1261 AAGAACAGCCAGGAGATACATATCTATGCAGAAAATGATGACGCCCATGTTACACAAA 1320
QY
1321 ATGTTACGCTGAATATAGAAAAGAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380
Db
1321 ATGTTACGCTGAATATAGAAAAGAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380
QY
1381 GCGTCCCTGTTCTGTAGTGGCTACCCGCTACCCCTTGGACCTGGAAGAAGTGTTCGGAC 1440
Db
1381 GCGTCCCTGTTCTGTAGTGGCTACCCGCTACCCCTTGGACCTGGAAGAAGTGTTCGGAC 1440
QY
1441 AAATCTCCCAATTCGACGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
|||||

Db
1441 AAATCTCCCAATTCGACGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
QY
1501 AAAGTGTTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATATGATGAGGCGGGAAAGGG 1560
Db
1501 AAAGTGTTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATATGATGAGGCGGGAAAGGG 1560
QY
1561 CTCTCGTCAAAATGCTGTGCGTACAATTTCTATGGGCACGCTTTCGCAAAACCATCTTTTA 1620
Db
1561 CTCTCGTCAAAATGCTGTGCGTACAATTTCTATGGGCACGCTTTCGCAAAACCATCTTTTA 1620
QY
1621 AACTACACAGGCCCTTCCCTTTTCATCCAAAGACAATCTCTCTATGCGACCATGTTGG 1680
Db
1621 AACTACACAGGCCCTTCCCTTTTCATCCAAAGACAATCTCTCTATGCGACCATGTTGG 1680
QY
1681 CTCTGCTCCCTTCATTTGTTCTCATGTGTTCTGATCTGCCACAAATACAAAAGCAA 1740
Db
1681 CTCTGCTCCCTTCATTTGTTCTCATGTGTTCTGATCTGCCACAAATACAAAAGCAA 1740
QY
1741 TTTAGTAGCAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTTGGATAACGAGTAC 1800
Db
1741 TTTAGTAGCAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTTGGATAACGAGTAC 1800
QY
1801 TTCTAGTTGACTTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCTCGAGAGAGAA 1860
Db
1801 TTCTAGTTGACTTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCTCGAGAGAGAA 1860
QY
1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGGGCTTTTCGGGAGGTGATGAAGCCACGCC 1920
Db
1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGGGCTTTTCGGGAGGTGATGAAGCCACGCC 1920
QY
1921 TATGGCATTTAGTAAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980
Db
1921 TATGGCATTTAGTAAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980
QY
1981 GCTGACAGTGTGAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
Db
1981 GCTGACAGTGTGAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
QY
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAAGAGAGAGAAAG 2100
Db
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAAGAGAGAGAAAG 2100
QY
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCCTCAACTTACCTAAGAGTAAAGAGAGAA 2160
Db
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCCTCAACTTACCTAAGAGTAAAGAGAGAA 2160
QY
2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATTAATTTTACCCTTACTTTTC 2220
Db
2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATTAATTTTACCCTTACTTTTC 2220
QY
2221 CAGGACATTTCAAAATTCAGCATGCTGTTACAGAGAGTTCAGTTTACACCCCTTTC 2280
Db
2221 CAGGACATTTCAAAATTCAGCATGCTGTTACAGAGAGTTCAGTTTACACCCCTTTC 2280
QY
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTTCAATTTTGAAGATGAGATGATGAA 2340
Db
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTTCAATTTTGAAGATGAGATGATGAA 2340
QY
2341 AACCAGAGAGGCTGGCAGAAAGAGAGAGATTTTGAACCTGTCTGAGTTTGAAGAC 2400
Db
2341 AACCAGAGAGGCTGGCAGAAAGAGAGAGATTTTGAACCTGTCTGAGTTTGAAGAC 2400
QY
2401 CTCTCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTCGAGTTCAAGTCTGT 2460
Db
2401 CTCTCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTCGAGTTCAAGTCTGT 2460
QY
2461 GTCCACAGAGACCTGGCAGCCAGGATGTTGGTCAACCCACGCGGAGGTGTTGAAGATC 2520
Db
2461 GTCCACAGAGACCTGGCAGCCAGGATGTTGGTCAACCCACGCGGAGGTGTTGAAGATC 2520
QY
2521 TGTGACTTTGGACTGGCCCGAGACATCTCCAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580
Db
2521 TGTGACTTTGGACTGGCCCGAGACATCTCCAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580


```
QY 2581 GCACGGCTGCGGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2581 GCACGGCTGCGGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640
QY 2641 AAGAGTGACGCTGCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGGAAC 2700
Db 2641 AAGAGTGACGCTGCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGGAAC 2700
QY 2701 CCTTACCTGGCATCTCTGTCAGCGTAACTTCTATAAAGTATTTTCACTGGATTTAA 2760
Db 2701 CCTTACCTGGCATCTCTGTCAGCGTAACTTCTATAAAGTATTTTCACTGGATTTAA 2760
QY 2761 ATGGAGCAGCCATCTATCCACAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2820
Db 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTCTTCCCACTGACTTCAATTTTAGGATGTGAGTGT 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTCTTCCCACTGACTTCAATTTTAGGATGTGAGTGT 2880
QY 2881 GCAGAGGAGAGAGCATGTATCCACAGAGCATCTATCCATCTACCAAAACAGGCGCCCT 2940
Db 2881 GCAGAGGAGAGAGCATGTATCCACAGAGCATCTATCCATCTACCAAAACAGGCGCCCT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGCGCAGGTGAAGATTACACAGAA 3000
Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGCGCAGGTGAAGATTACACAGAA 3000
QY 3001 AGAAGTTACGAGAGGCGCTTGAGCCCGCCAGCTAGAGGCTGTAGACCGCAGAGCA 3060
Db 3001 AGAAGTTAGCAGAGGCGCTTGAGCCCGCCAGCTAGAGGCTGTAGACCGCAGAGCA 3060
QY 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGCTGTCTGCTGAGCTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGCTGTCTGCTGAGCTTTCTCT 3120
QY 3121 AGATGCTGTCTGCCATTTACTCAAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTGTCTGCCATTTACTCAAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
QY 3181 GCGGAGAGCCATAATAGACTTGTGTGAGCGCGCTACAGCGCTGTCTGCTGAGCTTTCTCT 3240
Db 3181 GCGGAGAGCCATAATAGACTTGTGTGAGCGCGCTACAGCGCTGTCTGCTGAGCTTTCTCT 3240
QY 3241 AGCTTCAAGGAGAGCCATCTATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300
Db 3241 AGCTTCAAGGAGAGCCATCTATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTTGCTAAGGAAAGCTATAATGATTTTTAAAAATCTATGTTTAA 3360
Db 3301 ACCAAACCCGTTTTTTGCTAAGGAAAGCTATAATGATTTTTAAAAATCTATGTTTAA 3360
QY 3361 AATACATGTAATCTTTTCACTATTTAGTATATATTTTATGATGGAATAAATCTTC 3420
Db 3361 AATACATGTAATCTTTTCACTATTTAGTATATATTTTATGATGGAATAAATCTTC 3420
QY 3421 TACTGTAAAAA 3453
Db 3421 TACTGTAAAAA 3453
```

```
RESULT 8
PCR-US92-02750-1
; Sequence 1, Application PC/TUS9202750
; GENERAL INFORMATION:
; APPLICANT: LEMISCHKA, IHOR R.
; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell
; TITLE OF INVENTION: Receptors And Their Ligands
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
```

```
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: US
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02750
; FILING DATE: 19920402
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FEIT, IRVING N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..3009
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 31..3006
; PCT-US92-02750-1
```

```
Query Match 100.0%; Score 3453; DB 5; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCTGCTGCTACCGCGCGCTCCGGAGGCGCATCGGGCGTTGGCGCAGCGCAGCACCGG 60
Db 1 GCGGCTGCTGCTACCGCGCGCTCCGGAGGCGCATCGGGCGTTGGCGCAGCGCAGCACCGG 60
QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 CCATCATCTGCTACCGAATGCTGCGAGGATCCCGAGAGACCTTCCAGTGTACCCCGAGCGC 240
Db 181 CCATCATCTGCTACCGAATGCTGCGAGGATCCCGAGAGACCTTCCAGTGTACCCCGAGCGC 240
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGAGCGGTATATGAAGCGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 AGCTCCCTGGCTGCCAGCGCGCACTTTGATTACAAAACAGAGAAATCGTTTCCATGGCC 420
Db 361 AGCTCCCTGGCTGCCAGCGCGCACTTTGATTACAAAACAGAGAAATCGTTTCCATGGCC 420
QY 421 ATCTTGAACGTGACAGAGACCCAGCGAGAGAAATACCTACTCCATATTTAGAGCGAAACGC 480
Db 421 ATCTTGAACGTGACAGAGACCCAGCGAGAGAAATACCTACTCCATATTTAGAGCGAAACGC 480
QY 481 GCCAACTACACAGTACTGTTTCCAGTGAATGTAAGAGATACACAGTGTATGCTGTAAGG 540
Db 481 GCCAACTACACAGTACTGTTTCCAGTGAATGTAAGAGATACACAGTGTATGCTGTAAGG 540
```

Db 481 GCCAACTACACAGTACTGTTTACACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540
Qy 541 AGACCTTACTTTAGGAGATGGAACACAGAGTGCATCTCTGCTCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAGATGGAACACAGAGTGCATCTCTGCTCTCCGAGGGTGT 600
Qy 601 CCGAGGCCACTCTGGAGTGGGTGCTCTGAGCTGCCACAGGAAAGCTGTAAAGAGAA 660
Db 601 CCGAGGCCACTCTGGAGTGGGTGCTCTGAGCTGCCACAGGAAAGCTGTAAAGAGAA 660
Qy 661 GGCCCTGCTCTGTCAGAAAGGAGGAAAGTACTTATGAGTTGTCGGAACAGACATC 720
Db 661 GGCCCTGCTCTGTCAGAAAGGAGGAAAGTACTTATGAGTTGTCGGAACAGACATC 720
Qy 721 AGATGCTGCTAGAAATGACATGGGCGCGGAATGACCAAGCTGTTACCATAGATCTA 780
Db 721 AGATGCTGCTAGAAATGACATGGGCGCGGAATGACCAAGCTGTTACCATAGATCTA 780
Qy 781 AACCAGGCTCCTCAGAGCACATGCCCCAGTATTCTGAAAGTGGGGAAACCTTCTGG 840
Db 781 AACCAGGCTCCTCAGAGCACATGCCCCAGTATTCTGAAAGTGGGGAAACCTTCTGG 840
Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACTGGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGAGAGGAGGACCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960
Db 901 AAAGCCCTGAGAGGAGGACCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960
Qy 961 ATTTCGGATCTCTTGGGCTTTGCTCTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
Db 961 ATTTCGGATCTCTTGGGCTTTGCTCTCCGTTGGGAAGAACGACACCGGATATTACACC 1020
Qy 1021 TGCTCTTCTCAAAGCACCCAGCCAGTACGCGTTGGTGACCATCTTAGAAAAAGGGTTT 1080
Db 1021 TGCTCTTCTCAAAGCACCCAGCCAGTACGCGTTGGTGACCATCTTAGAAAAAGGGTTT 1080
Qy 1081 ATAAAGCTACAGCTCGGAAGAGATGTAATTTGACCCGTACGAAAGTTCGCTTC 1140
Db 1081 ATAAAGCTACAGCTCGGAAGAGATGTAATTTGACCCGTACGAAAGTTCGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGACGTTGGATCTCTCAAGCCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGACGTTGGATCTCTCTCAAGCCCTCA 1200
Qy 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
Db 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
Qy 1261 AAGAACAGCCAGAGAGTACATATCTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320
Db 1261 AAGAACAGCCAGAGAGTACATATCTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320
Qy 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTACCAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTACCAATGCTCAGCCAGCCAG 1380
Qy 1381 GCGTCCTGTTCTGATGGCTACCCGCTACCCCTTTGGACCTGGAAGAGTGTTCGAC 1440
Db 1381 GCGTCCTGTTCTGATGGCTACCCGCTACCCCTTTGGACCTGGAAGAGTGTTCGAC 1440
Qy 1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGAGTGTGGAATAAAGAGGCTTAACAGA 1500
Db 1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGAGTGTGGAATAAAGAGGCTTAACAGA 1500
Qy 1501 AAAGTGTTCGCCAGTGGGTGTCAGAGTACTCTAAATATGAGTGGAGGCGGGAAGGG 1560
Db 1501 AAAGTGTTCGCCAGTGGGTGTCAGAGTACTCTAAATATGAGTGGAGGCGGGAAGGG 1560
Qy 1561 CTTCTGCTCAAACTCTGCGTACAATTTCTATGGCAGCTCTTCGGAACCATCTTTTAA 1620
Db 1561 CTTCTGCTCAAACTCTGCGTACAATTTCTATGGCAGCTCTTCGGAACCATCTTTTAA 1620

Qy 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGAACATCTCTTCTATGACCACTTGG 1680
Db 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGAACATCTCTTCTATGACCACTTGG 1680
Qy 1681 CTCTGTCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Qy 1741 TTTAGGTACGAGAGTACGTCAGATGATCAGGTGACTGCCCCCTCGATTAACAGTAC 1800
Db 1741 TTTAGGTACGAGAGTACGTCAGATGATCAGGTGACTGCCCCCTCGATTAACAGTAC 1800
Qy 1801 TTCTACGTTGACTTTCAGGACTATGAATATCACCCTTAAGTGGGAGTTCGCCAGAGAGAC 1860
Db 1801 TTCTACGTTGACTTTCAGGACTATGAATATGACCTTAAGTGGGAGTTCGCCAGAGAGAC 1860
Qy 1861 TTAGAGTTTGGGAAGGCTCTGGGCTCTGGGCTTTTCGGGAGGGTGAAGACCCAGGCC 1920
Db 1861 TTAGAGTTTGGGAAGGCTCTGGGCTCTGGGCTTTTCGGGAGGGTGAAGACCCAGGCC 1920
Qy 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATCTTAAAGAGAAA 1980
Db 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATCTTAAAGAGAAA 1980
Qy 1981 GCTCACAGCTGTGAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTCACAGCTGTGAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACAACTCGTAATCTGCTGGGCGATGACACACTGTACAGGCCAGTGTACTTG 2100
Db 2041 CACCATGACAACTCGTAATCTGCTGGGCGATGACACACTGTACAGGCCAGTGTACTTG 2100
Qy 2101 ATTTTGAATATTTGCTATGTTGCTATGTTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAG 2160
Db 2101 ATTTTGAATATTTGCTATGTTGCTATGTTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAG 2160
Qy 2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAAACATATTTTCACTTCCCTACTTTC 2220
Db 2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAAACATATTTTCACTTCCCTACTTTC 2220
Qy 2221 CAGCACATTCAAATTCAGCATGCTGTTTCAGGAGAGTTCAGTTACACCCGCTTTC 2280
Db 2221 CAGCACATTCAAATTCAGCATGCTGTTTCAGGAGAGTTCAGTTACACCCGCTTTC 2280
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
Qy 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTTGAACGTCTGACGTTTGAAGAC 2400
Db 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTTGAACGTCTGACGTTTGAAGAC 2400
Qy 2401 CTCTTTGCTTTCGTTACCAAGTGGCCAAAGCATGGAATTCCTGAGTTCAAGTCGT 2460
Db 2401 CTCTTTGCTTTCGTTACCAAGTGGCCAAAGCATGGAATTCCTGAGTTCAAGTCGT 2460
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTTACCCACGGGAAGTGTGAAGAT 2520
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTTACCCACGGGAAGTGTGAAGAT 2520
Qy 2521 TGTGACTTTGACTGGCCCCGAGACATCCTGAGCGACTCCAGTACTGTCAGGCGCAAC 2580
Db 2521 TGTGACTTTGACTGGCCCCGAGACATCCTGAGCGACTCCAGTACTGTCAGGCGCAAC 2580
Qy 2581 GCAGGCTGCGCGTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2640
Db 2581 GCAGGCTGCGCGTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2640
Qy 2641 AAGAGTACGCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTTGGGTGTGAAC 2700
Db 2641 AAGAGTACGCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTTGGGTGTGAAC 2700

```
QY 2701 CCTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAACTGATTCAGAGTGGATTAA 2760
Db 2701 CCTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAACTGATTCAGAGTGGATTAA 2760
QY 2761 ATGAGCAGCCATTCATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
Db 2761 ATGAGCAGCCATTCATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
QY 2821 TTTGACTCAAGAGAGCGCCATCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGAGAGCGCCATCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2880
QY 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGCCCT 2940
Db 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGCCCT 2940
QY 2941 CAGCAGAGCGGGCTCAGAGCCAGTCGGCCACAGCGCCAGGTGAAGATTACAGAGAA 3000
Db 2941 CAGCAGAGCGGGCTCAGAGCCAGTCGGCCACAGCGCCAGGTGAAGATTACAGAGAA 3000
QY 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCACCTAGCAGCTGTAGACCGCAGAGCA 3060
Db 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCACCTAGCAGCTGTAGACCGCAGAGCA 3060
QY 3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCGCTTGTCTGCTGACTTTCTCT 3120
Db 3061 AGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCGCTTGTCTGCTGACTTTCTCT 3120
QY 3121 AGATGCTCTGCCATTTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTCTGCCATTTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
QY 3181 CGCGAGAGCAATAATCAGACTTGTGTGAGCGCCCTACCCCTGGGGGCTTTCCACG 3240
Db 3181 CGCGAGAGCAATAATCAGACTTGTGTGAGCGCCCTACCCCTGGGGGCTTTCCACG 3240
QY 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATAGTATATCTTGTAAATACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATAGTATATCTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGATGAAATAAACTTTC 3420
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGATGAAATAAACTTTC 3420
QY 3421 TACTGTAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAA 3453
Db 3421 TACTGTAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAA 3453
```

RESULT 9

PCT-US92-05401-1

; Sequence 1, Application PC/TUS9205401

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED

; STREET: 180 VARICK STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05401
; FILING DATE: 19920626
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31...3009
; NAME/KEY: mat_peptide
; LOCATION: 31...3006
; PCT-US92-05401-1
```

```
Query Match 100.0%; Score 3453; DB 5; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGCTGGCTACCGCGGCTCCGGAGGCGCATCGGGCGTTTGGCGAGCGAGCGACCGG 60
Db 1 CGGGCTGGCTACCGCGGCTCCGGAGGCGCATCGGGCGTTTGGCGAGCGAGCGACCGG 60
QY 61 CGGCTGCTGCTGCTTGTGTTTGTGCAAGTAATGTTTGTGACACCGTTTCAACACCAAG 120
Db 61 CGGCTGCTGCTGCTTGTGTTTGTGCAAGTAATGTTTGTGACACCGTTTCAACACCAAG 120
QY 121 CTGCGCTGTGATCAAGTGTGTTTAACTAGTATGAGAACAAATGGCTCATCAGCGGAAAG 180
Db 121 CTGCGCTGTGATCAAGTGTGTTTAACTAGTATGAGAACAAATGGCTCATCAGCGGAAAG 180
QY 181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTATCCCGGAGCGC 240
Db 181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTATCCCGGAGCGC 240
QY 241 CAGAGTGAAGGACGGTATATGAAGGGCCACCGTGGAGGTGGCCAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGACGGTATATGAAGGGCCACCGTGGAGGTGGCCAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTGCTCTGCTGCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTGCTCTGCTGCTTTAAGCAC 360
QY 361 AGCTCCCTGGGCTGCCAGCGCCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 361 AGCTCCCTGGGCTGCCAGCGCCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
QY 421 ATCTTTGAAGTGCAGAGACCGCAGGAGGAATACCTACTCATATTCAGAGCGAACGC 480
Db 421 ATCTTTGAAGTGCAGAGACCGCAGGAGGAATACCTACTCATATTCAGAGCGAACGC 480
QY 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
Db 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
QY 541 AGACCTTACTTTAGGAAGATGAAAACACAGGATGCACTCTCTGCACTCTCCGAGGCTTT 600
Db 541 AGACCTTACTTTAGGAAGATGAAAACACAGGATGCACTCTCTGCACTCTCCGAGGCTTT 600
QY 601 CCGAGAGCCACTGTGGAGTGGGTGCTGTCAGCTCCACAGGAAAGCTGTAAAGAA 660
Db 601 CCGAGAGCCACTGTGGAGTGGGTGCTGTCAGCTCCACAGGAAAGCTGTAAAGAA 660
```

Qy	661	GGCCCTGCTGTGTGCAGAAAGAGGAAAGGTACTTTCATGAGTTCTGTCGGAAACACATC	720	Qy	1741	TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCGGTACACGAGTAC	1800
Db	661	GGCCCTGCTGTGTGCAGAAAGAGGAAAGGTACTTTCATGAGTTCTGTCGGAAACACATC	720	Db	1741	TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCGGTACACGAGTAC	1800
Qy	721	AGATGCTGTGTAGAAATGCACTGGCCGCGGAAATGCAACCAAGCTGTTCCACCATAGATCTA	780	Qy	1801	TTTACGTTGACTTTCAGGAGCTATGAATATGACCTTAACTGGAGTTCCCGAGAGAAC	1860
Db	721	AGATGCTGTGTAGAAATGCACTGGCCGCGGAAATGCAACCAAGCTGTTCCACCATAGATCTA	780	Db	1801	TTTACGTTGACTTTCAGGAGCTATGAATATGACCTTAACTGGAGTTCCCGAGAGAAC	1860
Qy	781	AACCAAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGCG	840	Qy	1861	TTAGAGTTTGGAAAGGTCCTGGGCTGCGGCTTTCGGAGAGGTGATGACGCCACGGCC	1920
Db	781	AACCAAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGCG	840	Db	1861	TTAGAGTTTGGAAAGGTCCTGGGCTGCGGCTTTCGGAGAGGTGATGACGCCACGGCC	1920
Qy	841	ATCAGGTTGAAGGCATCATGTGAACCATGGATTCCGGCTCACCTGGAGCTGGAAGAC	900	Qy	1921	TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGCGGTGAAGATGCTTAAAGAGAAA	1980
Db	841	ATCAGGTTGAAGGCATCATGTGAACCATGGATTCCGGCTCACCTGGAGCTGGAAGAC	900	Db	1921	TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGCGGTGAAGATGCTTAAAGAGAAA	1980
Qy	901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGGACCATG	960	Qy	1981	GCTCAGAGCTGTGAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGGACCATG	960	Db	1981	GCTCAGAGCTGTGAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA	2040
Qy	961	ATTGGGATTCCTTGGCCCTTTGTCTTCCGTGGGAAGAACGACACCGGATATTACACC	1020	Qy	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG	2100
Db	961	ATTGGGATTCCTTGGCCCTTTGTCTTCCGTGGGAAGAACGACACCGGATATTACACC	1020	Db	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG	2100
Qy	1021	TGCTCTTCTCAAGACACCCAGCAGTCAGCGTTGGTGACCATCCTAGAAAAAGGGTTT	1080	Qy	2101	ATTTTGTGAATATTTGCTATGCTGACCTCCCTCAACTACTTAAGAAAGTAAAGAGAGAA	2160
Db	1021	TGCTCTTCTCAAGACACCCAGCAGTCAGCGTTGGTGACCATCCTAGAAAAAGGGTTT	1080	Db	2101	ATTTTGTGAATATTTGCTATGCTGACCTCCCTCAACTACTTAAGAAAGTAAAGAGAGAA	2160
Qy	1081	ATAAACGCTACAGCTCGCAAGAAAGATGATGAAATTTGACCGCTACGAAAGTTCTGCTTC	1140	Qy	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCCTACTTTC	2220
Db	1081	ATAAACGCTACAGCTCGCAAGAAAGATGATGAAATTTGACCGCTACGAAAGTTCTGCTTC	1140	Db	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCCTACTTTC	2220
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCAGCAATTCGATGCGACGTGATCTTCTCAAGCCCTCA	1200	Qy	2221	CAGGCACATTCAAATTCAGCATGCTGCTTCAAGTTCAGTTCACACCCGCGCTTG	2280
Db	1141	TCAGTCAGGTTTAAAGCGTACCCAGCAATTCGATGCGACGTGATCTTCTCAAGCCCTCA	1200	Db	2221	CAGGCACATTCAAATTCAGCATGCTGCTTCAAGTTCAGTTCACACCCGCGCTTG	2280
Qy	1201	TTTCTCTGTGAACAGAGGCTCGAGGATGGGTACACATATCTTAAATTTTGGCATCAT	1260	Qy	2281	GATCAGCTCTCAGGTTCAATGGGAAATTCATTCATTCATTCATTCATTCATTCATTCAT	2340
Db	1201	TTTCTCTGTGAACAGAGGCTCGAGGATGGGTACACATATCTTAAATTTTGGCATCAT	1260	Db	2281	GATCAGCTCTCAGGTTCAATGGGAAATTCATTCATTCATTCATTCATTCATTCATTCAT	2340
Qy	1261	AAGAACAGCGAGAGAGTACATATTCATGCAGAAATATGATGACCGCCAGTTTACACAAA	1320	Qy	2341	AACCAAGAGAGGCTGCGCAGAAAGAGAGGAGGAGGATTTTGAACGCTGCTGAGCTTGAAG	2400
Db	1261	AAGAACAGCGAGAGAGTACATATTCATGCAGAAATATGATGACCGCCAGTTTACACAAA	1320	Db	2341	AACCAAGAGAGGCTGCGCAGAAAGAGAGGAGGAGGATTTTGAACGCTGCTGAGCTTGAAG	2400
Qy	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380	Qy	2401	CTCCTTCTGCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAGTCTGT	2460
Db	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380	Db	2401	CTCCTTCTGCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAGTCTGT	2460
Qy	1381	CGGTCTCTTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC	1440	Qy	2461	GTCCACAGAGACTGCGCAGCAGGAATGTGTTGTCACCCACCGGAGGTTGTTGAAGATC	2520
Db	1381	CGGTCTCTTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC	1440	Db	2461	GTCCACAGAGACTGCGCAGCAGGAATGTGTTGTCACCCACCGGAGGTTGTTGAAGATC	2520
Qy	1441	AAATCTCCCAATTTGCACGAGGAAATCCAGAAAGGTTTGAATATAAGAGGCTAACAGA	1500	Qy	2521	TGTGACTTTGGACTGGCCCGAGACATCCTCAGCGACTCCAGCTACGTCGTCAGGGGCAAC	2580
Db	1441	AAATCTCCCAATTTGCACGAGGAAATCCAGAAAGGTTTGAATATAAGAGGCTAACAGA	1500	Db	2521	TGTGACTTTGGACTGGCCCGAGACATCCTCAGCGACTCCAGCTACGTCGTCAGGGGCAAC	2580
Qy	1501	AAAGTGTGTCAGTGGGTGCGAGAGTACTCTTAAATATGATGAGCGCCGGGAAAGG	1560	Qy	2581	GCACGGCTGCCGGTGAAGTGGATGGCCCGGAGAGCTTATTTGAAGGGATCTACAAATC	2640
Db	1501	AAAGTGTGTCAGTGGGTGCGAGAGTACTCTTAAATATGATGAGCGCCGGGAAAGG	1560	Db	2581	GCACGGCTGCCGGTGAAGTGGATGGCCCGGAGAGCTTATTTGAAGGGATCTACAAATC	2640
Qy	1561	CTTCTGGTCAATGCTGTGCTACAAATTCATGGCAGCTCTTGGCAACCATCTTTTAA	1620	Qy	2641	AAGAGTGAGTCTGCTGCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Db	1561	CTTCTGGTCAATGCTGTGCTACAAATTCATGGCAGCTCTTGGCAACCATCTTTTAA	1620	Db	2641	AAGAGTGAGTCTGCTGCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Qy	1621	AACTCACCAGGCCCTTCCCTTTTCAATCAAGACACATCTCTCTTCTATGCGACCATGGG	1680	Qy	2701	CTTACCTCTGGCATTCCTGTCGACGCTAACTTCTTAACTGATTCAGAGTGGATTTAAA	2760
Db	1621	AACTCACCAGGCCCTTCCCTTTTCAATCAAGACACATCTCTCTTCTATGCGACCATGGG	1680	Db	2701	CTTACCTCTGGCATTCCTGTCGACGCTAACTTCTTAACTGATTCAGAGTGGATTTAAA	2760
Qy	1681	CTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA	1740	Qy	2761	ATGGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT	2820
Db	1681	CTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA	1740	Db	2761	ATGGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT	2820
				Qy	2821	TTTGACTCAAGAGAGCGGCCATCTTCCCAACCTCCTTCATTTTAGGATGTCAGCTG	2880

Db 2821 TTTGACTCAAGGAGCGCCATCTTCCCAACCTGACTTCATTTTAGGATGTACGCTG 2880
QY 2881 CGAGAGGAGAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940
Db 2881 CGAGAGGAGAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940
QY 2941 CAGCAGAGAGGCGGCTCAGAGCGCCAGTCGCCACAGCGCCAGTGAAGATTCACAGAGAA 3000
Db 2941 CAGCAGAGAGGCGGCTCAGAGCGCCAGTCGCCACAGCGCCAGTGAAGATTCACAGAGAA 3000
QY 3001 AGAAGTTAGCAGAGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGACAGGCA 3060
Db 3001 AGAAGTTAGCAGAGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGACAGGCA 3060
QY 3061 AGATTAGCCTGCCCTCTGAGGAGCGCCCTACAGCGGCTTGTTCGCTGGACCTTTCTCT 3120
Db 3061 AGATTAGCCTGCCCTCTGAGGAGCGCCCTACAGCGGCTTGTTCGCTGGACCTTTCTCT 3120
QY 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAAACCTCTCTCGCACAG 3180
Db 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAAACCTCTCTCGCACAG 3180
QY 3181 CGCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCGCTACCTCGGGGCGCTTTCCAGG 3240
Db 3181 CGCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCGCTACCTCGGGGCGCTTTCCAGG 3240
QY 3241 AGCTTGAGGAGGAAAGCCATGTATCTGAATATAGTATATCTTTGTAATACGTGAACAA 3300
Db 3241 AGCTTGAGGAGGAAAGCCATGTATCTGAATATAGTATATCTTTGTAATACGTGAACAA 3300
QY 3301 ACCAAACCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATAATCTATGTTTAA 3360
Db 3301 ACCAAACCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATAATCTATGTTTAA 3360
QY 3361 AATACTATGTAATCTTTTTCATCTATTAGTATATATTTATGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAATCTTTTTCATCTATTAGTATATATTTATGATGGAATAAATCTTC 3420
QY 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

RESULT 10
PCT-US92-09893-1
; Sequence 1, Application PC/TUS9209893
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09893
; FILING DATE: 19921116
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7PT
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
PCT-US92-09893-1
Query Match 100.0%; Score 3453; DB 5; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGGCGCTGCTACCGCGCGCTCGGAGGCGCATCGGGCGTTGGCGGAGCGCAGCGACCGG 60
Db 1 GCGGCGCTGCTACCGCGCGCTCGGAGGCGCATCGGGCGTTGGCGGAGCGCAGCGACCGG 60
QY 61 CGGCTCTGCTGCTCTTCTTTGTCAGTAATCATTTCTTGAGACCGTTTACAAACCAAGAC 120
Db 61 CGGCTCTGCTGCTCTTCTTTGTCAGTAATCATTTCTTGAGACCGTTTACAAACCAAGAC 120
QY 121 CTGCGCTGTGATCAAGTGTGTTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAG 180
Db 121 CTGCGCTGTGATCAAGTGTGTTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAG 180
QY 181 CCATCATCTGCTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Db 181 CCATCATCTGCTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTCTGCTCTGGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTCTGCTCTGGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGCTGCGAGCGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGCC 420
Db 361 AGCTCCCTGGCTGCGAGCGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGCC 420
QY 421 ATCTTGAAGTGCAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGCAACGC 480
Db 421 ATCTTGAAGTGCAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGCAACGC 480
QY 481 GCCAATACACAGTACTGTTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAATACACAGTACTGTTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540
QY 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCATCTCGGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCATCTCGGAGGGTGT 600
QY 601 CCGGAGCCCACTGTGAGTGGTGTCTGCTGAGTCCACAGGGAAGCTGTAAAGAGAA 660
Db 601 CCGGAGCCCACTGTGAGTGGTGTCTGCTGAGTCCACAGGGAAGCTGTAAAGAGAA 660
QY 661 GCGCCTGCTGTTGTCAGAAAGGAGGAAAGTACTTTCATGAGTTGTTGCGGAACAGATC 720
Db 661 GCGCCTGCTGTTGTTTCTGAGAAAGGAGGAAAGTACTTTCATGAGTTGTTGCGGAACAGATC 720

Db 661 GGCCTGCTGTTGTAGAAAGGAGAAAGGTACTTTCATGAGTTGTTCCGAAACAGACATC 720
Qy 721 AGATGCTGTGCTAGAAATGCACTGGCCGCGGAATGCACCAAGCTGTTCCACCATAGATCTA 780
Db 721 AGATGCTGTGCTAGAAATGCACTGGCCGCGGAATGCACCAAGCTGTTCCACCATAGATCTA 780
Qy 781 AACCAGGCTCCTCAGAGCAGACTGCCCCAGTTATTTCTCTGAAAGTGGGGAACCCCTTGTGG 840
Db 781 AACCAGGCTCCTCAGAGCAGACTGCCCCAGTTATTTCTCTGAAAGTGGGGAACCCCTTGTGG 840
Qy 841 ATCAGGTGTAGGCCATTCATGTGAACCATGGATTCGGGCTACCTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAGGCCATTCATGTGAACCATGGATTCGGGCTACCTGGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGGACCATG 960
Db 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGGACCATG 960
Qy 961 ATTCCGATTTCTTTGGCCTTTGTGTTCTTCCGTGGGAAGGACGACCCGGATATTACACC 1020
Db 961 ATTCCGATTTCTTTGGCCTTTGTGTTCTTCCGTGGGAAGGACGACCCGGATATTACACC 1020
Qy 1021 TGCTTTCTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCATCTTAGAAAAAGGTTT 1080
Db 1021 TGCTTTCTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCATCTTAGAAAAAGGTTT 1080
Qy 1081 ATAAACCTACCACTCCAGAGAGAGTATCAAAATTGACCCGTACGAAAGATTTCTGCTTC 1140
Db 1081 ATAAACCTACCACTCCAGAGAGAGTATCAAAATTGACCCGTACGAAAGATTTCTGCTTC 1140
Qy 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCAGTGGATTTCTCTCAAGCCCTCA 1200
Db 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCAGTGGATTTCTCTCAAGCCCTCA 1200
Qy 1201 TTTCCCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Db 1201 TTTCCCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Qy 1261 AAGAACAGCCAGGAGAGTACATATTTATCCAGAAATGATGACGCCAGCTTACCCAAA 1320
Db 1261 AAGAACAGCCAGGAGAGTACATATTTATCCAGAAATGATGACGCCAGCTTACCCAAA 1320
Qy 1321 ATGTTCAAGCTGAAATAAGAAAGAAACCTCAAGTGCATAGCAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTCAAGCTGAAATAAGAAAGAAACCTCAAGTGCATAGCAATGCTCAGCCAGCCAG 1380
Qy 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440
Db 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440
Qy 1441 AAATCTCCCAATTCGACGGAGGAATCCAGAAAGAGTTTGAATTAAGAAAGCTTAACAGA 1500
Db 1441 AAATCTCCCAATTCGACGGAGGAATCCAGAAAGAGTTTGAATTAAGAAAGCTTAACAGA 1500
Qy 1501 AAAGTGTGTCAGTGGGTGAGGAGTACTCTAAATATGAGTGAGCCGCGGAAGGG 1560
Db 1501 AAAGTGTGTCAGTGGGTGAGGAGTACTCTAAATATGAGTGAGCCGCGGAAGGG 1560
Qy 1561 CTCTGTGTCAAATGCTGTGCTACAAATCTATGGGACGCTTTCGAAACCATCTTTTAA 1620
Db 1561 CTCTGTGTCAAATGCTGTGCTACAAATCTATGGGACGCTTTCGAAACCATCTTTTAA 1620
Qy 1621 AACTCAGAGCCCTTCCCTTTTCATCCAGAACACATCTCTCTATGACACCATTTGGG 1680
Db 1621 AACTCAGAGCCCTTCCCTTTTCATCCAGAACACATCTCTCTATGACACCATTTGGG 1680
Qy 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCCAATATCAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCCAATATCAAAAGCAA 1740
Qy 1741 TTTAGGTACGAGAGTCAGCTCAGATGATCAGGTGACTGCCCCCTTGATTAACAGATAC 1800
Db 1741 TTTAGGTACGAGAGTCAGCTCAGATGATCAGGTGACTGCCCCCTTGATTAACAGATAC 1800

Qy 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860
Qy 1861 TTAGAGTTTGGGAAGGTCCTGGGCTTGGCGTTTGGGAGGTTGATGAACGCCACGGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCCTGGGCTTGGCGTTTGGGAGGTTGATGAACGCCACGGCC 1920
Qy 1921 TATGCAATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1921 TATGCAATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Qy 1981 GCTGACGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTACAGGGCAGTGTACTTG 2100
Db 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTACAGGGCAGTGTACTTG 2100
Qy 2101 ATTTTGAATATTTGCTATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAG 2160
Db 2101 ATTTTGAATATTTGCTATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAG 2160
Qy 2161 TTTACAGAGATGCACAGAGATTTTAAAGAACATAATTTTACAGTCTTACCCACTTTC 2220
Db 2161 TTTACAGAGATGCACAGAGATTTTAAAGAACATAATTTTACAGTCTTACCCACTTTC 2220
Qy 2221 CAGGCACATTTCAAATTCAGCATGCTGTTCCAGCAGAAAGTTCAGTTACACCCGCCCTTG 2280
Db 2221 CAGGCACATTTCAAATTCAGCATGCTGTTCCAGCAGAAAGTTCAGTTACACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATGAA 2340
Qy 2341 AACACAGAGAGCTGCGAGAGAGAGAGAGATTTTCAACGCTGCTGACGTTTGAAGAC 2400
Db 2341 AACACAGAGAGCTGCGAGAGAGAGAGAGATTTTCAACGCTGCTGACGTTTGAAGAC 2400
Qy 2401 CTCTTTGCTTTGCTTACCAAGTGGCATGGAATTTCTGGAGTTTCAAGTCTGT 2460
Db 2401 CTCTTTGCTTTGCTTACCAAGTGGCATGGAATTTCTGGAGTTTCAAGTCTGT 2460
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTCACCCAGGGAAGTGTGGAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTCACCCAGGGAAGTGTGGAAGATC 2520
Qy 2521 TGTGACTTTGACTGCCCCGAGACATCCTGAGGGACTCCAGCTACGTCGAGGGGCAAC 2580
Db 2521 TGTGACTTTGACTGCCCCGAGACATCCTGAGGGACTCCAGCTACGTCGAGGGGCAAC 2580
Qy 2581 GCACGGCTGCCGCTCAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Db 2581 GCACGGCTGCCGCTCAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Qy 2641 AAGAGTACGCTGTGCTCTACGGCATCTCTCTCTGGAGATATTTTACTGGGTGTGAAC 2700
Db 2641 AAGAGTACGCTGTGCTCTACGGCATCTCTCTCTGGAGATATTTTACTGGGTGTGAAC 2700
Qy 2701 CCTTACCCCTGGATTCCTGTCAGCGTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2760
Db 2701 CCTTACCCCTGGATTCCTGTCAGCGTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2760
Qy 2761 ATGAGGACGACATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGCT 2820
Db 2761 ATGAGGACGACATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTTTCTAGATGTACGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTTTCTAGATGTACGCTG 2880

2392	Db	AACCAAGAAGAGCTGCGAAGAAGAAGAGAGATTGAACTGCTGACGTTGAAGAC	2415
2401	QY	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTCGAGTTCAAAGTCGTGT	2460
2452	Db	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTCGAGTTCAAAGTCGTGT	2511
2461	QY	GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCACCCACGGGAAGTGGTGAAGATC	2520
2512	Db	GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCACCCACGGGAAGTGGTGAAGATC	2571
2521	QY	TGTGACTTTGGACTGGCCCGAGACATCCTCGAGCGACTCCAGTACGTCTGACGGGCAAC	2580
2572	Db	TGTGACTTTGGACTGGCCCGAGACATCCTCGAGCGACTCCAGTACGTCTGACGGGCAAC	2631
2581	QY	GCACGGCTCCGGTGAAGTGGATGGCACC CGAGAGCTTATTGAAGGATCTACACAATC	2640
2632	Db	GCACGGCTCCGGTGAAGTGGATGGCACC CGAGAGCTTATTGAAGGATCTACACAATC	2691
2641	QY	AAGAGTGACGTCTGGTCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
2692	Db	AAGAGTGACGTCTGGTCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2751
2701	QY	CTTTACCTTGGCATTCCTGTGACGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA	2760
2752	Db	CTTTACCTTGGCATTCCTGTGACGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA	2811
2761	QY	ATGGACAGCCATTCTATGCCACAGAAGGATATACTTTTAATGCAATCCTGCTGGGCT	2820
2812	Db	ATGGACAGCCATTCTATGCCACAGAAGGATATACTTTTAATGCAATCCTGCTGGGCT	2871
2821	QY	TTTGACTCAAGGAAGCGGCATCCTTCCCAAACCTGACTTCATTTTAGATGTCTAGCTG	2880
2872	Db	TTTGACTCAAGGAAGCGGCATCCTTCCCAAACCTGACTTCATTTTAGATGTCTAGCTG	2931
2881	QY	GCAGAGCGAAGAAGC-----ATGTTATCAGAAACATCCATCC	2917
2932	Db	GCAGAGCGAAGAAGCGATTATCAGACATCGGTGGCAACGTCCAGCAACATCCATCC	2991
2918	QY	ATCTACCAAAACAGCGGCCCTCAGCAGAGAGCGGGCTCAGAGCC-CAGTCGCCACAG	2976
2992	Db	ATCTACCAAAACAGCGGCCCTCAGCAGAGAGCGGGCTCAGAGCGGCCATCGCCACAG	3051
2977	QY	CGCCAGGTGAAGATTACAGAGAAAGAAAGTTACGGAGGAGGCCTTGACCCCGCCACCCT	3036
3052	Db	GCCAGGTGAAGATTACAGAGAAAGAAAGTTACGGAGGAGGCCTTGACCCCGCCACCCT	3111
3037	QY	AGCAGGCTGAGACCGCAGACCAAGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCG	3096
3112	Db	AGCAGGCTGAGACCGCAGACCAAGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCG	3171
3097	QY	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCCATTACTTCCAAAGTGACTTCAT	3156
3172	Db	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCCATTACTTCCAAAGTGACTTCAT	3231
3157	QY	AAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATAATGAGACTTTGTTGGTAGCCC	3216
3232	Db	AAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATAATGAGACTTTGTTGGTAGCCC	3291
3217	QY	GCCTACCTCGGGGCCCTTT-----CCACGAGCTTGAGGGGAAGCCATGATCTGAA	3268
3292	Db	GCCTACCTCGGGGGGCTTTCCAGGCCGCCAGCTTGAGGGGAAGCCATGATCTGAA	3351
3269	QY	ATATAGTATATCTTGTAAATACGTGAACAAACAAACCCGTTTTTTCGTAAGGGGAAG	3328
3352	Db	ATATAGTATATCTTGTAAATACGTGAACAAACAAACCCGTTTTTTCGTAAGGGGAAG	3411
3329	QY	CTAAATATGATTTTAAAAATCTATGTTTTAAAAATCTATGTAACTTTTTCATCTATTTA	3388
3412	Db	CTAAATATGATTTTAAAAATCTATGTTTTAAAAATCTATGTAACTTTTTCATCTATTTA	3471
3389	QY	GTGATATATTTATGGATGGAATAAACTTTTCTACTGTAAAAAATAAAAAA	3438
3472	Db	GTGATATATTTATGGATGGAATAAACTTTTCTACTGTAAAAAATAAAAAA	3521

Db 352 ACCCTGCAAGTGCAGCTGCCACCCAGGAGACCTTCTCGCTCTGGGCTCTTTAAGCAC 411
Qy 361 AGCTCCCTGGGCTGCCACGCCGACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 412 AGCTCCCTGGGCTGCCACGCCGACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 471
Qy 421 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCCATATTTCAGAGCGAAGCC 480
Db 472 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCCATATTTCAGAGCGAAGCC 531
Qy 481 CCCAACTACAGTACTGTTCCAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540
Db 532 GCCAACTACAGTACTGTTCCAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 591
Qy 541 AGACCTTACTTTAGGAAGATGAAACCCAGAGTGCACCTGCTGTCATCTCCGAGGGTGT 600
Db 592 AGACCTTACTTTAGGAAGATGAAACCCAGAGTGCACCTGCTGTCATCTCCGAGGGTGT 651
Qy 601 CCGAGGCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAGAAGAA 660
Db 652 CCGAGGCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAGAAGAA 711
Qy 661 GGCCCTGCTGTGTGTCAGAAAGAGGAAAAGTACTTTCATGAGTTGTTCCGAACAGACATC 720
Db 712 GGCCCTGCTGTGTGTCAGAAAGAGGAAAAGTACTTTCATGAGTTGTTCCGAACAGACATC 771
Qy 721 AGATGCTGTGTAGAAATGCACCTGGGCGCGAATGCAACAGCTGTTCCACATAGATCTA 780
Db 772 AGATGCTGTGTAGAAATGCACCTGGGCGCGAATGCAACAGCTGTTCCACATAGATCTA 831
Qy 781 AACCAGGCTCCTCAGAGCACATGCCCCAGTTATTCTTGAAGTGGGGAAACCTTTGG 840
Db 832 AACCAGGCTCCTCAGAGCACATGCCCCAGTTATTCTTGAAGTGGGGAAACCTTTGG 891
Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC 900
Db 892 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC 951
Qy 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTACCAAAACAGGACCATG 960
Db 952 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTACCAAAACAGGACCATG 1011
Qy 961 ATTCGGATTCCTTTGGGCTTTGTGCTCCGTGGGAAGAACGACACCGGATATTACACC 1020
Db 1012 ATTCGGATTCCTTTGGGCTTTGTGCTCCGTGGGAAGAACGACACCGGATATTACACC 1071
Qy 1021 TGCTCTTCCCTCAAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080
Db 1072 TGCTCTTCCCTCAAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1131
Qy 1081 ATAAACGCTACAGCTCCCAAGAGAGTATGAATTTGACCCGTAGCAAAAGTTCTGCTTC 1140
Db 1132 ATAAACGCTACAGCTCCCAAGAGAGTATGAATTTGACCCGTAGCAAAAGTTCTGCTTC 1191
Qy 1141 TCAGTCAGGTTTAAAGCTACCCAGCAATCCGATCGACGTGATCTCTCAAGCCCTCA 1200
Db 1192 TCAGTCAGGTTTAAAGCTACCCAGCAATCCGATCGACGTGATCTCTCAAGCCCTCA 1251
Qy 1201 TTTCTTCTGTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTTAAATTTTTCGATCAT 1260
Db 1252 TTTCTTCTGTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTTAAATTTTTCGATCAT 1311
Qy 1261 AAGAACAGCCAGAGAGTACATATTCTATCGAGAAATGATGACGCCCGAGTTTACACAAA 1320
Db 1312 AAGAACAGCCAGAGAGTACATATTCTATCGAGAAATGATGACGCCCGAGTTTACACAAA 1371
Qy 1321 ATGTTACGCTGTAATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Db 1372 ATGTTACGCTGTAATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1431
Qy 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTGAGCTGGAAGAGTGTTCGGAC 1440
|||||

Db 1432 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTGTTGACCTGGAAGAGTGTTCGGAC 1491
Qy 1441 AAATCTCCCAATTTGCACGGAGAAATCCAGAGGAGTTTGGAAATAAAGGCTACACAGA 1500
Db 1492 AAATCTCCCAATTTGCACGGAGAAATCCAGAGGAGTTTGGAAATAAAGGCTACACAGA 1551
Qy 1501 AAAGTGTTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGAAGCCGGAAGGG 1560
Db 1552 AAAGTGTTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGAAGCCGGAAGGG 1611
Qy 1561 CTTCTGGTCAAAATGCTGTCGCTACAAATCTATGCGGCACAGTCTTCCGAAACCATCTTTTA 1620
Db 1612 CTTCTGGTCAAAATGCTGTCGCTACAAATCTATGCGGCACGTTCTTCCGAAACCATCTTTTA 1671
Qy 1621 AACTCACAGAGCCCTTCCCTTTTCATCAAGACAACATCTCTTCTATGCGACCATTTGGG 1680
Db 1672 AACTCACAGAGCCCTTCCCTTTTCATCAAGACAACATCTCTTCTATGCGACCATTTGGG 1731
Qy 1681 CTTCTGCTCCCTTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1740
Db 1732 CTTCTGCTCCCTTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1791
Qy 1741 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACGAGTAC 1800
Db 1792 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACGAGTAC 1851
Qy 1801 TTTCTAGGTGACTTTCAGGACTATGAATATGACCTTAAGTGGAGTTTCCGAGAGAGAAC 1860
Db 1852 TTTCTAGGTGACTTTCAGGACTATGAATATGACCTTAAGTGGAGTTTCCGAGAGAGAAC 1911
Qy 1861 TTTAGAGTTTGGGAAGGCTCGGGGCTCGGGCTTTTCGGAGGGTGTGAACGCTCACGGCC 1920
Db 1912 TTTAGAGTTTGGGAAGGCTCGGGGCTCGGGCTTTTCGGAGGGTGTGAACGCTCACGGCC 1971
Qy 1921 TATGCAATTAGTAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
Db 1972 TATGCAATTAGTAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 2031
Qy 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 2032 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2091
Qy 2041 CACCATTGACAAATCGTGAATCTGCTGGGGATGCACACTGTCCAGGGCCAGTGTACTTG 2100
Db 2092 CACCATTGACAAATCGTGAATCTGCTGGGGATGCACACTGTCCAGGGCCAGTGTACTTG 2151
Qy 2101 ATTTTGAATATGTTGCTATGTTGACCTCCTCACTACTTAAGAGTAAAAGAGAGAAG 2160
Db 2152 ATTTTGAATATGTTGCTATGTTGACCTCCTCACTACTTAAGAGTAAAAGAGAGAAG 2211
Qy 2161 TTTTCACAGGACATGGACAGATTTTAAAGACATTAATTTTCACTTACCCCTACTTTC 2220
Db 2212 TTTTCACAGGACATGGACAGATTTTAAAGACATTAATTTTCACTTACCCCTACTTTC 2271
Qy 2221 CAGGCACATTTCAAAATCCAGCATGCTGCTGTTACAGAGAAGTTCAGTTACACCCGCTTG 2280
Db 2272 CAGGCACATTTCAAAATCCAGCATGCTGCTGTTACAGAGAAGTTCAGTTACACCCGCTTG 2331
Qy 2281 GATCAGCTCTCAGGTTCAATGGAAATCAATTCATTCTGAAGATGAGATTGAATATGAA 2340
Db 2332 GATCAGCTCTCAGGTTCAATGGAAATCAATTCATTCTGAAGATGAGATTGAATATGAA 2391
Qy 2341 AACCAGAGAGGCTGGCAGAAAGAGGAGGAGAGATTTGAACGTGCTGACGTTTGAGAC 2400
Db 2392 AACCAGAGAGGCTGGCAGAAAGAGGAGGAGAGATTTGAACGTGCTGACGTTTGAGAC 2451
Qy 2401 CTCTTTTGTCTTTCGCTTACCAAGTGGCCAAAGCATGSAATTCCTGAGTTCAAGTCGTGT 2460
Db 2452 CTCTTTTGTCTTTCGCTTACCAAGTGGCCAAAGCATGSAATTCCTGAGTTCAAGTCGTGT 2511
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAGATGTGTTGGTCACCCACGGGAGGTGGTGAAGATC 2520
Db 2512 GTCCACAGAGACCTGGCAGCAGGAGATGTGTTGGTCACCCACGGGAGGTGGTGAAGATC 2571

Db 532 GCCAACTACACAGTACTGTTCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 591
Qy
541 AGACCTTACTTTAGNAGATGAAACACAGGATGACTGCTCTGCATCTCCAGGGGTGT 600
Db
592 AGACCTTACTTTAGNAGATGAAACACAGGATGACTGCTCTGCATCTCCAGGGGTGT 651
Qy
601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAGAAGAA 660
Db
652 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAGAAGAA 711
Qy
661 GGCCTGTCTGTGTCAGAAAGAGGAAAGGTACTTATGATGTGTCGGAACAGACATC 720
Db
712 GGCCTGTCTGTGTCAGAAAGAGGAAAGGTACTTATGATGTGTCGGAACAGACATC 771
Qy
721 AGATGCTGTCTAGAAATGCACTGGCCGCGGAATGCACCAAGCTGTTACCATAGATCTA 780
Db
772 AGATGCTGTCTAGAAATGCACTGGCCGCGGAATGCACCAAGCTGTTACCATAGATCTA 831
Qy
781 AACCAGGCTCCTCAGAGCACACTGCCCAAGTTATTCTGAAAGTGGGGAAACCTTGTGG 840
Db
832 AACCAGGCTCCTCAGAGCACACTGCCCAAGTTATTCTGAAAGTGGGGAAACCTTGTGG 891
Qy
841 ATCAGGTGAAGCCATCCATGTGAACCATGGATTGGGCTCACCTGGGAGCTGGAAGAC 900
Db
892 ATCAGGTGAAGCCATCCATGTGAACCATGGATTGGGCTCACCTGGGAGCTGGAAGAC 951
Qy
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATG 960
Db
952 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATG 1011
Qy
961 ATTCGGATTCTCTGGCCCTTTGTGCTTCCGTTGGGAAGCAACGCGGATATTACACC 1020
Db
1012 ATTCGGATTCTCTGGCCCTTTGTGCTTCCGTTGGGAAGCAACGCGGATATTACACC 1071
Qy
1021 TGCTCTTCTCTCAAGCACCCAGCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080
Db
1072 TGCTCTTCTCTCAAGCACCCAGCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1131
Qy
1081 ATAACGCTTACCAGTCCGAAAGACAGATGAAATTGACCCGTACGAAAGTTCTGCTTC 1140
Db
1132 ATAACGCTTACCAGTCCGAAAGACAGATGAAATTGACCCGTACGAAAGTTCTGCTTC 1191
Qy
1141 TCAGTCAGGTTTAAAGCGTACCACAGAAATCCGATGCAGTGGATCTTCTCAAGCCTCA 1200
Db
1192 TCAGTCAGGTTTAAAGCGTACCACAGAAATCCGATGCAGTGGATCTTCTCAAGCCTCA 1251
Qy
1201 TTTCTCTGTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Db
1252 TTTCTCTGTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311
Qy
1261 AAGAACAGCCAGGAGAGTACATATTCTATGCAGAAATGATGACGCCAGTTTCAACCAA 1320
Db
1312 AAGAACAGCCAGGAGAGTACATATTCTATGCAGAAATGATGACGCCAGTTTCAACCAA 1371
Qy
1321 ATGTTCCAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380
Db
1372 ATGTTCCAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1431
Qy
1381 GCGTCTCTGTCTGTATGGGTACCCGCTACCCCTCTTGAGCCTGGAAGAAAGTTCTCGGAC 1440
Db
1432 GCGTCTCTGTCTGTATGGGTACCCGCTACCCCTCTTGAGCCTGGAAGAAAGTTCTCGGAC 1491
Qy
1441 AAATCTCCCAATTGCACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAGGCTTAACAGA 1500
Db
1492 AAATCTCCCAATTGCACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAGGCTTAACAGA 1551
Qy
1501 AAAGTGTGTCAGGAGGAGTCTGAGCAGTACTCTAAATATGACTGAGCCCGGGGAAGG 1560
Db
1552 AAAGTGTGTCAGGAGGAGTCTGAGCAGTACTCTAAATATGAGTGAAGCCGGGAAGG 1611
Qy
1561 CTTCCTGGTCAATGCTGTGCGTACAAATTTATGCGGCACGCTGTGCGAAACCACTTTTFA 1620
Db
1612 CTTCCTGGTCAATGCTGTGCGTACAAATTTATGCGGCACGCTGTGCGAAACCACTTTTFA 1671

Qy 1621 AACTCACCAGGCCCTTCCTTTTCATCCAAAGACAACATCTCCTTCTATGCGACCATTTGG 1680
Db
1672 AACTCACCAGGCCCTTCCTTTTCATCCAAAGACAACATCTCCTTCTATGCGACCATTTGG 1731
Qy
1681 CTCTGTCTCCCTTTCATTTGTTTCTATTGTGTTGATCTGCCACAATAACAAAAAGCAA 1740
Db
1732 CTCTGTCTCCCTTTCATTTGTTTCTATTGTGTTGATCTGCCACAATAACAAAAAGCAA 1791
Qy
1741 TTTAGTAGCAGAGTCAAGTGCAGATGATCCAGTGACTGGCCCCCTCGGATACAGAGTAC 1800
Db
1792 TTTAGTAGCAGAGTCAAGTGCAGATGATCCAGTGACTGGCCCCCTCGGATACAGAGTAC 1851
Qy
1801 TTCTAGTTGACTTCAAGGACTATGAATATGACTTAAGTGGGAGTTCCCGAGAGAGAAC 1860
Db
1852 TTCTAGTTGACTTCAAGGACTATGAATATGACTTAAGTGGGAGTTCCCGAGAGAGAAC 1911
Qy
1861 TTAGAGTTTGGGAAGTCTCGGCTTTCGGGAGGTTGATGAACGCCACGCC 1920
Db
1912 TTAGAGTTTGGGAAGTCTCGGCTTTCGGGAGGTTGATGAACGCCACGCC 1971
Qy
1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db
1972 TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 2031
Qy
1981 GCTGACAGCTGCARAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
Db
2032 GCTGACAGCTGCARAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2091
Qy
2041 CACCATGACAACATCGTGAATCTGCTGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2100
Db
2092 CACCATGACAACATCGTGAATCTGCTGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2151
Qy
2101 ATTTTGAATATGTTGCTATGGTGACCTCCTCAACTACCTAAGAGATAAAGAGAGAAG 2160
Db
2152 ATTTTGAATATGTTGCTATGGTGACCTCCTCAACTACCTAAGAGATAAAGAGAGAAG 2211
Qy
2161 TTTTCAGACACATGGACAGAGATTTTAAAGACATAAATTCAGTCTTACCCCTACTTTC 2220
Db
2212 TTTTCAGACACATGGACAGAGATTTTAAAGACATAAATTCAGTCTTACCCCTACTTTC 2271
Qy
2221 CAGGCACATTCAAATTCACAGCATGCCTGGTTACAGAGAAGTTTCAGTTACACCCGCCCTTG 2280
Db
2272 CAGGCACATTCAAATTCACAGCATGCCTGGTTACAGAGAAGTTTCAGTTACACCCGCCCTTG 2331
Qy
2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATCTGAAGATGAGATTGAATATGAA 2340
Db
2332 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATCTGAAGATGAGATTGAATATGAA 2391
Qy
2341 AACCAGAGAGGCTGGCAGAAAGAGAGGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400
Db
2392 AACCAGAGAGGCTGGCAGAAAGAGAGGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2451
Qy
2401 CTCTTTTGTGTTGCGTACCAAGTGGCAGAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
Db
2452 CTCTTTTGTGTTGCGTACCAAGTGGCAGAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2511
Qy
2461 GTCCACAGAGACTGSCAGCCAGGAATGTGTTGGTCAACCCACGGGAAGTGGTGAAGATC 2520
Db
2512 GTCCACAGAGACTGSCAGCCAGGAATGTGTTGGTCAACCCACGGGAAGTGGTGAAGATC 2571
Qy
2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGACTCAGCTACGTCGTCAGGGGCAAC 2580
Db
2572 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGACTCAGCTACGTCGTCAGGGGCAAC 2631
Qy
2581 GCACGGCTCCCGTGAAGTGGATGGCACCCGAGAGCTTATTGAAGGATCTACACAATC 2640
Db
2632 GCACGGCTCCCGTGAAGTGGATGGCACCCGAGAGCTTATTGAAGGATCTACACAATC 2691
Qy
2641 AAGAGTGAGCTGTGGTCTTACGGCATCTCTCTGCGGAGATATTTTCACTGGGTGAAC 2700
Db
2692 AAGAGTGAGCTGTGGTCTTACGGCATCTCTCTGCGGAGATATTTTCACTGGGTGAAC 2751

2701 CCTTACCTGGCATTCTCTGCACGCTAACTTCTATAAATGATTCAGAGTGATTTAA 2760
2752 CTTTACCTGGCATTCTCTGCACGCTAACTTCTATAAATGATTCAGAGTGATTTAA 2811
2761 ATGGAGCAGCCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCCTGCGGCT 2820
2812 ATGGAGCAGCCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCCTGCGGCT 2871
2821 TTTGACTCAAGGAAGGCGCATCTTCCCAACCTGACTTCAATTTAGGATGTCAGTG 2880
2872 TTTGACTCAAGGAAGGCGCATCTTCCCAACCTGACTTCAATTTAGGATGTCAGTG 2931
2881 GCAGAGGAGAGAAAGC-----ATGTATCAGAACATCATCC 2917
2932 GCAGAGGAGAGAAAGCGATGTATCAGAACATGGGTGGCAAGTCCCAGAACATCATCC 2991
2918 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGGGGCTCAGAGCC-CAGTGGCCACAG 2976
2992 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGGGGCTCAGAGCGCCCATCGCCACAG 3051
2977 GCCAGGTGAAGATTCACAGAGAAAGAGTTAGCGAGGAGGCTTGGACCCCGCCACCT 3036
3052 GCCCAGGTGAAGATTCACAGAGAAAGAGTTAGCGAGGAGGCTTGGACCCCGCCACCT 3111
3037 AGCAGGCTGTAGACGCGCAGCAAGATTAGCTCGCCCTCGAGGAAGCGCCCTCAGAGC 3096
3112 AGCAGGCTGTAGACGCGCAGCAAGATTAGCTCGCCCTCGAGGAAGCGCCCTCAGAGC 3171
3097 CGTTCCTGCTGGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGAATCTAT 3156
3172 CGTTCCTGCTGGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGAATCTAT 3231
3157 AAAATCAAACTCTCTCCACAGCGGGGAGAGCAATATAGAGCTTTGTTGGTAGCCC 3216
3232 AAAATCAAACTCTCTCCACAGCGGGGAGAGCAATATAGAGCTTTGTTGGTAGCCC 3291
3217 GCCTACCTGGGGGCTTT-----CCACGAGCTTGAGGGAAGCCATGTATCTGAA 3268
3292 GCCTACCTGGGGGCTTTCCAGGCCCCCGAGGCTTGAGGGAAGCCATGTATCTGAA 3351
3269 ATATAGTATATCTTTGTAATACGTGAACAAACCAACCCGTTTTTTCTAAGGGAAG 3328
3352 ATATAGTATATCTTTGTAATACGTGAACAAACCAACCCGTTTTTTCTAAGGGAAG 3411
3329 CTAATATATATTTTAAATCTATGTTTTTAAATACTATGTAACCTTTTCTATCTATTA 3388
3412 CTAATATATATTTTAAATCTATGTTTTTAAATACTATGTAACCTTTTCTATCTATTA 3471
3389 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAA 3438
3472 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAA 3521

RESULT 14
US-07-977-451-3
; Sequence 3, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,451
FILING DATE: 19921119
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US UNASSIGNED
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
US-07-977-451-3

Query Match 64.5%; Score 2227.2; DB 1; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 8 GGCTACCGCGCTCCGGAGCCATCGGCGGCTTGGCGGACGCGAGCGGCGGCTGC 67
DB 35 GGGGACCCCGGCTCCGGAGCCATCGGCGGCTTGGC---GCGGACGCGGCGGCGGCTGC 91
QY 68 TGCTGCTGTTGTTTGTCTAGTAATGTTCTTGAGACCGTTACAAACCAAGACCTGCCTG 127
DB 92 CGCTGCTGTTGTTTGTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151
QY 128 TGATCAAGTGTGTTTAAATCAATGCTCATGAGAACCAATGCTCATGAGCGGGAAGCCATCAT 187

152 TGATCAAGTGTGTTTAAATCAATCATAAAGACAATGATTCATCAGTGGGGAAGTCATCAT 211
188 CGTACCCAAATGGTCCGAGGATCCCAAGAACCTCCAGTGTACCCGAGGCCGAGAGTG 247
212 CATATCCATGGTATCAGATCCCGGAAGACCTCGGGTGTGCGTTGAGACCCGAGAGCT 271
248 AAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCCATCACCCTGC 307
272 CAGGGACAGTGTACGAGCTGCCCTGTGGAGTGGATGTATCTGCTTCCATCACACTGC 331
308 AAGTGCAGCTGGCACCCGAGGAGACCTTTCTCCCTCTCGGCTTTAAGCACAGCTCCC 367
332 AAGTGCCTGGTGCATGCCCGAGGAACATTTCTCTCTGGCTCTTAAGCACAGCTCCC 391
368 TGGCTGCCAGCCGACTTTGATTTACAAACACAGGAATCGTTTCCATGCCATCTTGA 427
392 TGAATGCCAGCCACATTTTGATTTACAAAACAGAGGAGTTGTTTCCATGGTCTAATTTGA 451
428 ACGTGACAGAGACCCGAGGAGAAATACCTTACTCCATATTCAGAGGGAAGCGGCCAACT 487
452 AATGACAGAAACCCAGCTGGAGAAATACCTTACTTTTATTCAGAGTGAAGCTACCAAT 511
488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547
512 ACACATATTTGTTACAGTGTAGTATAAGAAATACCTTGTACACATTAAGAAGACCTT 571
548 ACTTAGAGAGATGGAAAACCCAGGATGCAGTGTCTGATCTCCGAGGTTCCGGAGC 607
572 ACTTAGAAAATGGAAAACCCAGGACCCCTGGTCTGCATATCTGAGAGCCTCCAGAGC 631
608 CCAGTGTGGAGTGGTGTCTGACGTCCCAAGGGAAGCTGTAAAGAAAGAGCCCTG 667
632 CGATCGTGGATGGGTGCTTGCATTCACAGGGGAAGCTGTAAAGAAAGTCCAG 691
668 CTGTTGTGAGAAAAGGAAAAGTACTTCTATGAGTTGTTCCGAACACATCAGATGCT 727
692 CTGTTGTTAAAAGGAGGAAAAGTCTTTCATGAATTTTGGGACGACATAGGTGCT 751
728 GTGCTAGAAATGCACTGGGCGGCAATGCCACCAAGCTGTTCCACCATAGATCTAAACAGG 787
752 GTGCAAGAAATGAACTGGGAGGAATGCCACAGGCTGTTCCACATAGATCTAAATCAA 811
788 CTCTCAGAGCACACTGCCCGAGTTATTCGAAAGTGGGGAACCTTGTGGATCAGGT 847
812 CTCTCAGACCAATGCCCAATTTTCTTAAGTAGGGGAACCTTTATGGATAAGGT 871
848 GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCCTGGAGCTGGAAGACAAAGCCC 907
872 GCAAGCTGTTCATGTGAACCATGATTCGGGCTCACCCTGGGAATTAAGAAACAAAGCAC 931
908 TGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATCGGA 967
932 TCGAGGAGGGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991
968 TTCTCTTGGCTTTGTCTTCCCTGGGAAGGAGACACCCGGATATTACACCTGCTCTT 1027
992 TTCTGTTGCTTTGTATCATCATGTCAGTGGCAAGAAACAGACCCGGATACTACACTTGTTCCT 1051
1028 CCTCAAGACCCCGAGCAGTCCAGGTTGGTGACCATCTAGAAAAGGTTTATAAAGC 1087
1052 CTTCAGACATCCCAAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGATTTATAATG 1111
1088 CTACCACTGCGAAGAGAGTATGAAATTTGACCCGTAGCAAAAGTCTCTCTCAGTCA 1147
1112 CTACCAATTCAGTGAAGATTATGAAATTTGACCAATATGAGAGTTTGTGTTTCTGTCA 1171
1148 GGTTTAAGGCTACCCAGGAATCCGATGCGAGTGGATCTCTCAAGCCCTCATTTCTCT 1207
1172 GGTTTAAGGCTACCCCAAAATCAGATGATAGTGGACCTTCTCTCGAAAATCATTTCTCT 1231
1208 GTGACAGAGAGGCTTGAGGATGGGTACACCATCTAAATTTTGGCATCATRAGACA 1267

1232 GTGAGCAAAAGGGTCTTGTATAACGGATACAGCATATCCAAGTTTGTCAATCATAAAGCAC 1291
1268 AGCCAGGAGACTACATATTTCTATCAGAAAATGATGAGCCCAAGTTCCACAAAATGTTCA 1327
1292 AGCCAGGAGAAATATATTTCCATGAGAAAATGATGATGCCCAATTTACAAAATGTTCA 1351
1328 CGCTGAATATAGAAAAGAACTCAAGTGTCTAGCAAAATGCCCTCAGCCAGCCAGGCTCCT 1387
1352 CGCTGAATATAGAAAGAACTCAAGTGTCTGCGAAGACATCGCGAAGTCAAGGCTCCT 1411
1388 GTTCTCTGATGTACCGTACCGTCTTGACCTTGAGACCTGGAAGAGTGTTCGGACAAATCTC 1447
1412 GTTCTCGATGGATACCCATTTACCATCTGGACCTGGAAGAGTGTTCAGACAAGTCTC 1471
1448 CCAATTCACAGGAGAAATCCAGAAAGAGTGTGAATTAAGAGCTTACAGAAAAGTCT 1507
1472 CCAACTGCACAGAGAGATCACAGAAAGAGTCTGGAATAGAAAAGCTTACAGAAAAGTCT 1531
1508 TTGGCCAGTGGTGTGAGCAGTACTCTAAATATGAGTGAGCCGGGAAAGGCTTCTCG 1567
1532 TTGGACAGTGGTGTGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGTTCCTCG 1591
1568 TCAATGCTGTGCGTACAAATTTCTATGGCAGCTCTTGGCAACCATCTTTTAAATCTAC 1627
1592 TCAAGTGTGTGCATACAAATTCCTTGGCAGATCTGTGAGACGATCTCTTTAAATCTCTC 1651
1628 CAGGCCCTTCCCTTTTCATCCAAGACACATCTCTCTATCGGACCATTTGGGCTCTGTC 1687
1652 CAGGCCCTTCCCTTTTCATCCAAGACACATCTCTCTATGCAACAATTTGGTGTGTC 1711
1688 TCCCTCTCATTTGTTGTTCTCATTTGATCTGCCACAATAACAAAAGCAATTTAGGT 1747
1712 TCCCTCTCATTTGCTTTTAAACCTGCTAATTTGTGCACAGTACAAAAGCAATTTAGGT 1771
1748 ACGAGAGTCACTGCAGATGATCCAGGTGACTTGGCCCTCGATGAACGAGTACTTCTACG 1807
1772 ATGAAAACGAGCTACAGATGTTGACAGTGACCGGCTCTCAGATAATGAGTACTTCTACG 1831
1808 TTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCGCGAGAGAGACTTAGAGT 1867
1832 TTGATTTTCAGAGAAATGAATATGATCAAAATGGGAGTTTCCAAGAGAAAATTTAGAGT 1891
1868 TTGGAAGTCTCTGGGCTTGGCGGAGTGTGATGAACGCGACGCTTATTTAGCA 1927
1892 TTGGAAGTGTACTAGGATCAGGTGCTTTTGGAAAAGTGTGAACGCAAGCTTATGGAA 1951
1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAAGAGAAAGCTGACA 1987
1952 TTAGCAAAACAGGAGTCTCAATCCAGTTGCCGTCAAAATGCTGAAAGAAAAGCAGACA 2011
1988 GCTGTGAAAAGAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
2012 GCTCTGAAAAGAGAGGAGCTCATGTGAGAACTCAAGATGATGATGATGATGATGATGATG 2071
2048 ACACATCTGCAATCTGCTGGGCGATGCACACTCTCAGGGCGAGTGTACTTGTGATTTTG 2107
2072 AGAATTTGTGAACCTGCTGGGCGGTGCACACTGTGAGAGCAATTTACTTGTGATTTTG 2131
2108 AATATTGTTGCTATGTTGACCTCTCAACTACTAAGAGTAAAAGAGAGAGTTCACA 2167
2132 AATACTGTTGCTATGTTGATCTCTCACTATCTAAGAGTAAAAGAGAGAAATTTTACA 2191
2168 GGACATGACAGAGATTTTAAAGAACATAATTTCAAGTCTTCTTCTTCTTCTTCTTCTTCT 2227
2192 GGACTTGCACAGAGATTTTCAAGGAACACAATTTCAAGTCTTCTTCTTCTTCTTCTTCTTCT 2251
2228 ATTCAAAATCCAGATGCTGTTTCAAGAGAGTTCAGTTACACCGGCTTGGATCAGC 2287
2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCGGCTTGGATCAAA 2311
2288 TCTCAGGCTTCAATGGGAATCAATTCATCTGAGAGTGAATTAATTAATTAATTAATTAAT 2347
2312 TCTCAGGCTTCAATGGGAATTCATTTCTCTGAGAGTGAATTAATTAATTAATTAATTAAT 2371

TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
US-08-252-517-3

Query Match 64.58; Score 2227.2; DB 1; Length 3501;

Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY	8	GGCTACCGCGCTCCGGAGGCCATCGCGGCTTGGCGCAGCGCAGCGACCGCGGCTGC	67
DB	35	GGGACCCCGGCTCCGGAGGCCATCGCGGCTTGGC---CGCGACGGGCAACGCTGC	91
QY	68	TGCTGCTGTGTTTGTCTAGTATGATCTCTGAGACCGTTACAAACCAAGACCTGCCTG	127
DB	92	CGCTGCTGCTGTTTCTGCAATGATATTGGGACTATTACAATCAAGATCGCCTG	151
QY	128	TGATCAAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGGAAGCCATCAT	187
DB	152	TGATCAAGTGTGTTTAAATCAATCATAGAACATGATTCATCAGTGGGGAAGTCATCAT	211
QY	188	CGTACCGAATGTTGCGAGGATCCCAAGACCTCCAGTGTACCCGAGCGCCAGAGTG	247
DB	212	CATATCCATGGTATCAGATCCCGGAAGACCTCGGGTGTGCTTGAGACCCAGAGCT	271
QY	248	AAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGTCCATCACCTGC	307
DB	272	CAGGACAGTGTACGAGCTGCCGCTGTGGAGTGGATGATCTGCTCCATCACACTGC	331
QY	308	AAGTGCAGTGGCCACCCAGGGGACCTTCTCTGCTCTGGTCTTTAAGCACAGCTCCC	367
DB	332	AAGTGTGCTGATGCTGCCAGGGAACATTTCTCTCTGCTCTGGTCTTTAAGCACAGCTCCC	391
QY	368	TGGCTGCCAGCGCCTTTGATTTACAAACAGAGAACTGTTTCCATGCCATCTTGA	427
DB	392	TGAATGCCAGCCACATTTTGATTTACAAACAGAGAGTGTGTTCCATGGCTATTGA	451
QY	428	ACGTGACAGACCGCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGCGCCCAACT	487
DB	452	AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAAT	511
QY	488	ACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT	547
DB	512	ACACAATATTGTTTACAGTGAGTATAGAATAACCTGCTTTACACATTAAGAAGACCTT	571
QY	548	ACTTTAGAGATGGAAAACCCAGATGACAGTCTGCTGCACTCCAGAGGTTTCCGGAGC	607
DB	572	ACTTTAGAAAATGGAAAACCCAGACGCCCTGGTCTGCAATATCTGAGAGCGTCCAGAGC	631
QY	608	CCACTGTGGAGTGGTCTCTGACGCTCCCAAGGAAAGCTGTAAAGAAAGAGCCCTG	667
DB	632	CGATCGTGGATGGTGTGCTTTCGATTCACAGGGGGAAGCTGTAAAGAAAGATCCAG	691
QY	668	CTGTTGTGAGAAAAGGAGAAAAGTACTTTCATGAGTTGTTCCGAAACACATCAGATGCT	727
DB	692	CTGTTGTAAAAGAGGAAAAGTGTCTCATGATATTGTTGGGACGACATAGGTGCT	751
QY	728	GTGCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTTCCATCAGATATCAACAGG	787
DB	752	GTGCCAGAAATGAATGGGAGGAATGCACCAAGCTGTTTCAAAATAGATCTAAATCAA	811
QY	788	CTCCTCAGACCACTGCCCCAGTTATCTCTGAAAGTGGGGGAACCCCTTGTGGATCAGGT	847

DB	812	CTCCTCAGACCACATTGCCACAAATATTCTTAAAGTAGGGGAACCCCTTATGATAAGGT	871
QY	848	GTAAGGCCATCCATGTTGAACCATGGATTCGGGCTCACCTGGAGCTGGAGACAAAGCCC	907
DB	872	GCAAAGCTGTTTCATGTTGAACCATGGATTCGGGCTCACCTGGAGCTGGAGACAAAGCAC	931
QY	908	TGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATGATTCGGA	967
DB	932	TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAACAGAACTATGATACGGA	991
QY	968	TTCTCTTGGCCTTTGTGCTTTCGTTGGAAAGAACGACACCGGATATTACACTGCTCTT	1027
DB	992	TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAACGACACCGGATACTACACTGTTCTCT	1051
QY	1028	CCTCAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTCAAAAAGGTTTATAAAGC	1087
DB	1052	CTTCAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTGTAAGAGGATTTATAAATG	1111
QY	1088	CTACAGCTCGCAAGAAGATGAAATTTGACCCGTACGAAAGTTCTCTCTCTCACTCA	1147
DB	1112	CTACCAATTCAGTGAAGATTATGAATTTGACCAATATGAAGATTTCTTTTCTGTCA	1171
QY	1148	GGTTTAAAGCTACCCACGAATTCGATGCAGTGGATCTTCTCAAGCCTCATTTCCCT	1207
DB	1172	GGTTTAAAGCTACCCACAAATCAGATGTACGTGGACCTTCTCTCGAAAATCATTTCTCT	1231
QY	1208	GTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTCCGATCATAGAACA	1267
DB	1232	GTGACAAAAGGGTCTTGATAACGGATACAGCATATCCAAGTTTTCGAATCAAGACCC	1291
QY	1268	AGCCAGGAGAGTACATATTTCTATCAGAAAATGATGACGCCAGTTACCAAAATTTTCA	1327
DB	1292	AGCCAGGAGATATATATTTCCATGCAAAAATGATGATGCCCAATTTACCAAAATTTCA	1351
QY	1328	CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCAGCCAGGCTCT	1387
DB	1352	CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGCAGAACATCGCGAAGTCAAGGCTCT	1411
QY	1388	GTTCTCTGTATGGCTACCCGCTACCCCTCTTGGACCTGGAAGTGTTCGGACAAATCTC	1447
DB	1412	GTTTCTGGATGGATACCCATACCATCTTGGACCTGGAAGTGTTCAGACAAAGTCTC	1471
QY	1448	CCATTCGACGAGGAATCCAGAGAGTGTGGAATAAAGGCTTAACAGAAAAGTGT	1507
DB	1472	CCAACTGCACAGAAGAGATCACAGAAGAGTCTGGAATAGAAGGCTTAACAGAAAAGTGT	1531
QY	1508	TTGGCCAGTGGTGTGAGCAGTACTCTAAATATGATGAGCGCGGAAAGGCTTCTTG	1567
DB	1532	TTGNACAGTGGTGTGAGCAGTACTCTAAACATGATGAGCCATAAAGGCTTCTTG	1591
QY	1568	TCAATGCTGTGCGTACAAATCTATGGCAGCTCTTGGAAACCACTCTTTTAAACTAC	1627
DB	1592	TCAAGTGTGTGCATACAAATCCCTTGGCACATCTTGTGAGACGATCTTTTAAACTCTC	1651
QY	1628	CAGCCCTTCCCTTTCATCCAAAGACATCTCTCTCTATCGACCATTTGGCTCTGTC	1687
DB	1652	CAGGCCCTTCCCTTTCATCCAAAGACATCTCTCTCTATCGACCATTTGGTGTGTC	1711
QY	1688	TCCCTCTCATTTGTTCTCTCTATGTCACACAAATACAAAAAGCAATTTAGGT	1747
DB	1712	TCTCTCTCATTTGCTTTTAAACCTGCTAATTTGTACAAAGTACAAAAGCAATTTAGGT	1771
QY	1748	ACGAGAGTCACTGCAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTACTTCTACG	1807
DB	1772	ATGAAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAGATAATGATGACTTCTACG	1831
QY	1808	TTGACTTCAGGCACTATGAATATGACCTTAAGTGGAGTTCCTCCGAGAGCACTTAGGT	1867
DB	1832	TTGATTTTCAGAGAATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGGT	1891
QY	1868	TTGGGAAGTCTCTGGGCTGCGCTTTTCGGGAGGTTGATGAACGCCACGGCTATGGCA	1927

1892 TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAAGTGATGAACGCAACAGCTTATGGAA 1951
1928 TTAGTAAACGGGAGTCTCAATTCAGCTGCGGTGAAGATGCTAAAGAGAGAAAGCTGCACA 1987
1952 TTAGCAAAACAGGAGTCTCAATTCAGGTTCGCGTCAAAATGCTGAAGAAAGAAAGCAGACA 2011
1988 GCTGTGAAAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
2012 GCTCTGAAGAGAGGCACTCATGTGCAAGTCAAGATGATGACCCAGCTGGGAAGCCAGC 2071
2048 ARAACATCGTAATCTGCTGGGGCATGCACACTGTCAGGCCAGTGTACTTGTATTTTG 2107
2072 AGAATATGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATTTACTTGTATTTTG 2131
2108 AATATTGCTGCTATGTGTGACCTCTCAACTACCTAAGAGTAAAGAGAGAGTTCACA 2167
2132 AATACTGTTGCTATGTGTGATCTCTCAACTATCTAAGAGTAAAGAGAGAAATTTTACA 2191
2168 GGACATGCACAGAGATTTTAAAGGAACATAATTTTCAAGTCTTACCCCTACTTTCCAGGCAC 2227
2192 GGACTTGCACAGAGATTTTCAAGGAACACAAATTTTCAAGTCTTACCCCTACTTTCCCAATCAC 2251
2228 ATTCAAAATCCAGATGCGCTGTTTACAGAGAGTTCAGTACACCCGCTTGGATCAGC 2287
2252 ATCCAAATTCAGCATGCGCTGTTTCAAGAGAGAGTTCAGATACACCCGCTTGGATCAAA 2311
2288 TCTCAGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTAATGAAATGAAACCCAGA 2347
2312 TCTCAGGCTTCATGGGAATTCATTTTCACTCTGAAGATGAAATTTTGAATGAAACCCAGA 2371
2348 AGAGCTGGGAGAGAGAGGAGAGATTTTGAACGCTGCTGACGCTTTTGAAGACCTCTCTTT 2407
2372 AAGGCTG-----GAAGAGAGGAGGACTTGAATGCTTACATTTGAAGATCTTCTTT 2425
2408 GCTTTGGCTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAGTCTGCTGTCACACA 2467
2426 GCTTTGCTATCAAAATGCTGCAAGAGATGGAATTTCTGGAATTTAAGTCTGCTGTTTACA 2485
2468 GAGACCTGGCAGCAGAGATGCTGTTGTCACCCAGGAGGTGCTGAAGATCTGTGACT 2527
2486 GAGACCTGGCCGCGCAGGAGACGCTGTTGTCACCCAGGAGAGTGGTGAAGATGCTGACT 2545
2528 TTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACGCTGCTCAGGGGCAACGCAAGGC 2587
2546 TTGGATTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTCAGGGCAATGCCCGTC 2605
2588 TGCCGGTGAAGTGAATGTCACCCGAGAGCTTATTTGAAGGATCTACACAAATCAAGAGTG 2647
2606 TGCCCTGTAATGGAATGGATGGCCCGCAAGAGCTGTTTGAAGGCACTTACACCAATTAAGAGTG 2665
2648 ACCTGCTGCTACGGCATCTCTCTGCGAGATATTTTCTGCGGTGCTGAACCTTACC 2707
2666 ATGCTGCTCATATGGAATATTTACTGCGGGAATCTTCTCAGTGGTGTGAATCTCTACC 2725
2708 CTGGCATTCCTGCTGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAATGGAGC 2767
2726 CTGGCATTCGGTGTGATGCTTAATCTACAACTGATTCAAATGGAATTTAAATGGATC 2785
2768 AGCCATTCATGCCACAGAGGATATATCTTTGTAATGCAATCTCTGCTGGGCTTTTGACT 2827
2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCTCTGCTGGGCTTTTGACT 2845
2828 CAAGGAAGGGCCATCTTCCCAACCTGACTTATTTTAGGATGTCAGCTGGGAGG 2887
2846 CAAGGAAGGGCCATCTTCCCTAAATTTGACTTCTGTTTGGATGTCAGCTGGGAGATG 2905
2888 CAGAAGAAG-----CATGTATCAGAACATCCATCCATCTACC 2924
2906 CAGAAGAAGCGATGATCAGATGCGGATGGCGGTGTTTCGGAATGCTCTCACACCTACC 2965
2925 AAAACAGGGCGGCTTCAAGCAGAGAGGCGGCTCAGAGC-CCAGTGGCCACAGCGCCAGG 2983
2966 AAAACAGGGCGGCTTTCAGCAGAGAGATGGAATTTGGGGCTACTCTCTCCGCGGCTCAGG 3025

3084 T-GAAGATTACAGAGAAAGAGTTAGCGAGGAGGCTTTGAGCCCGCC-----ACCCTAG 3038
3026 TCGAAGATTCTGACAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
3039 CAGCTGTAGACCCAGAGCCAGATAGCTCGCCTCT--GAGGAAGCGCCCTACACG 3096
3086 CAGCTGTAGATTACCAAAACAGATTAATTTTCACTAAAGAAATCTTATATCAAC 3145
3097 CGTTGCTTCGCTGGACTTTTCTTAGATGCTGCTGCCATTACTC-----CAAAGTCA 3149
3146 TGCTGCTTCACGACTTTTCTTAGAGCGCTGCGCTTACTCTTGTCTTTCAAAGGA 3205
3150 CTCTATAAAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATATGAGACTTGTGG 3209
3206 CTTTGTGAAATCAAAATCATCTCTGTCACAAGCGAGGAGCTGATAATGAACCTTTATTG 3265
3210 TGAGCCCGCTTACCTGGGGCTTTCCACGAGCTTGAGGGAAAGCCATGTATCTGAA 3269
3266 GAGCATTTGATCTGATCCCAAGGCTTCTCAGCGCGCTTGAGTGAATTTGTACCTGAAG 3325
3270 TATAGTATATTCTTGAATACGTGAAACAAACCAACCGCTTTTCTGTAAGGGAAGC 3329
3326 TACAGTATATTCTTGTAAATACATAAAACAAA-----AGCATTTTCTAGGAGAGC 3378
3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
3379 TAATATGATTTT--AAGTCTATCTTTTAAATAATATGTAAATTTTTCAGCTATTAG 3436
3390 TGATATATTTTATGATGGAATAAATCTTCTACTGTAAATAAAATAAAATAAAATAAA 3449
3437 TGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAATAAAATAAAATAAA 3496
3450 AAAA 3453
3497 AAAA 3500

Search completed: May 25, 2003, 12:20:02
Job time : 153.496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 11:51:23 ; Search time 398.232 Seconds
(without alignments)
11449.495 Million cell updates/sec

Title: US-09-919-408-1

Perfect score: 3453

Sequence: 1 GCGGCTGGCTACCGCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	10	US-09-919-408-1
2	3453	100.0	3453	10	US-09-872-136-1
3	2227.2	64.5	3501	10	US-09-919-408-3
4	2227.2	64.5	3501	10	US-09-872-136-3
5	2081.8	60.3	3120	10	US-09-982-610-22
6	466.6	13.5	645	9	US-09-796-692-7538
7	466.6	13.5	645	9	US-10-040-862-7538
8	456.4	13.2	610	9	US-09-796-692-8085
9	456.4	13.2	610	9	US-10-040-862-8085
10	454.6	13.2	605	9	US-09-796-692-8927
11	448.8	13.0	597	9	US-10-040-862-8927
12	448.8	13.0	597	9	US-09-796-692-7976
13	448.8	13.0	601	9	US-10-040-862-8273
14	448	13.0	601	9	US-09-796-692-8273
15	448	13.0	601	9	US-10-040-862-8273
16	448	13.0	608	9	US-09-796-692-8132
17	448	13.0	608	9	US-10-040-862-8132
18	447.2	13.0	597	9	US-09-796-692-9010
19	447.2	13.0	597	9	US-10-040-862-9010

20	440	12.7	605	9	US-09-796-692-8047	Sequence 8047, Ap
21	440	12.7	605	9	US-10-040-862-8047	Sequence 8047, Ap
22	439.8	12.7	589	9	US-09-796-692-8630	Sequence 8630, Ap
23	439.8	12.7	589	9	US-10-040-862-8630	Sequence 8630, Ap
24	424.4	12.3	594	9	US-09-796-692-8135	Sequence 8135, Ap
25	424.4	12.3	594	9	US-10-040-862-8135	Sequence 8135, Ap
26	423.2	12.3	564	9	US-09-796-692-9025	Sequence 9025, Ap
27	423.2	12.3	564	9	US-10-040-862-9025	Sequence 9025, Ap
28	416.8	12.1	556	9	US-09-796-692-7883	Sequence 7883, Ap
29	416.8	12.1	556	9	US-10-040-862-7883	Sequence 7883, Ap
30	407.2	11.8	550	9	US-09-796-692-8494	Sequence 8494, Ap
31	407.2	11.8	550	9	US-10-040-862-8494	Sequence 8494, Ap
32	399.8	11.6	539	9	US-09-796-692-8255	Sequence 8255, Ap
33	399.8	11.6	539	9	US-10-040-862-8255	Sequence 8255, Ap
34	386.8	11.2	510	9	US-09-796-692-8268	Sequence 8268, Ap
35	386.8	11.2	510	9	US-10-040-862-8268	Sequence 8268, Ap
36	360.2	10.4	497	9	US-09-796-692-8408	Sequence 8408, Ap
37	360.2	10.4	497	9	US-10-040-862-8408	Sequence 8408, Ap
38	357.4	10.4	3992	10	US-09-944-807-9	Sequence 9, Appli
39	350	10.1	606	9	US-09-796-692-8114	Sequence 8114, Ap
40	350	10.1	606	9	US-10-040-862-8114	Sequence 8114, Ap
41	342.2	9.9	508	9	US-09-796-692-8032	Sequence 8032, Ap
42	342.2	9.9	508	9	US-10-040-862-8032	Sequence 8032, Ap
43	329	9.5	5084	9	US-09-954-531-1383	Sequence 1383, Ap
44	329	9.5	5084	9	US-10-192-867-3	Sequence 3, Appli
45	329	9.5	5084	10	US-09-967-768A-277	Sequence 277, App

ALIGNMENTS

RESULT 1
US-09-919-408-1
; Sequence 1, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTOPOENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: InClone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666
FILING DATE: 02 APR 1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054

```

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
    LENGTH: 3453 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    FRAGMENT TYPE: N-terminal
    FEATURE:

```

```
FEATURE: NAME/KEY: mat_peptide
          LOCATION: 112..3006
FEATURE: NAME/KEY: sig_peptide
          LOCATION: 31..111
FEATURE:
```

NAME/KEY: CDS
LOCATION: 31.3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
9-408-1

US-09-919-408-1

Query Match	100.0%;	Score 3453;	DB 10;	Length 3453;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3453;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	CGCGCTGGCTACCGCGCGCTCCGGAGGCCATCGCGCGCTTGGCGCAGCGCAGCGACGG	60
DB	1	CGCGCTGGCTACCGCGCTCCGGAGGCCATCGCGCGCTTGGCGCAGCGCAGCGACGG	60
QY	61	CGGCTGCTGCTGCTTGTGTTTGTCTAGTAATGATCTTGAGACCGTTACAACCAAGAC	120
DB	61	CGGCTGCTGCTGCTTGTGTTTGTCTAGTAATGATCTTGAGACCGTTACAACCAAGAC	120
QY	121	CTGCCTGTGATCAAGTGTGTTTTAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	180
DB	121	CTGCCTGTGATCAAGTGTGTTTTAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	180
QY	181	CCATCATCGTACCGAAATGGTGGCAGGATCCCCAGAGACCTCCAGTGTACCCCGAGGCG	240
DB	181	CCATCATCGTACCGAAATGGTGGCAGGATCCCCAGAGACCTCCAGTGTACCCCGAGGCG	240
QY	241	CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC	300
DB	241	CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC	300
QY	301	ACCTCGAAGTGACGCTCGCCACCCAGGGGACCTTTCCTCCCTCGGTCTTTTAAGCAC	360
DB	301	ACCTCGAAGTGACGCTCGCCACCCAGGGGACCTTTCCTCCCTCGGTCTTTTAAGCAC	360
QY	361	AGTCCCTGGGCTGCCAGCGCACCTTGATTTACAAACACAGAAATCGTTTCCATGGCC	420
DB	361	AGTCCCTGGGCTGCCAGCGCACCTTGATTTACAAACACAGAAATCGTTTCCATGGCC	420
QY	421	ATCTTGAACGTGACAGACCCAGGACGAGAATACCTACTCCATATTCAGAGGCAAGCC	480
DB	421	ATCTTGAACGTGACAGACCCAGGACGAGAATACCTACTCCATATTCAGAGGCAAGCC	480
QY	481	GCCAACTACACAGTACTGTTCCACAGTGAATTAAGAGATACACAGCTGTATGTGCTAAGG	540
DB	481	GCCAACTACACAGTACTGTTCCACAGTGAATTAAGAGATACACAGCTGTATGTGCTAAGG	540
QY	541	AGACCTTACTTTAGGAAGATGGAACCCAGGATGCATGCTCTGCATCTCCAGGGGTGTT	600

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786

FILING DATE: <Unknown>

APPLICATION NUMBER: US/09/021,324

FILING DATE: <Unknown>

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3453 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 112..3006

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 31..111

FEATURE:

NAME/KEY: CDS

LOCATION: 31..3009

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-872-136-1

Query Match 100.0%; Score 3453; DB 10; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGGCGCTGGTACCGCGCGCTCGGAGGCATCGGGCGGTGGCGCAGCGCAGCGCGG	60
DB	1	GGGCGCTGGTACCGCGCGCTCGGAGGCATCGGGCGGTGGCGCAGCGCAGCGCGG	60
QY	61	CGGCTCGTCTGCTGTTGTTGTCAGTAATGATTTCTGAGACCGGTTACAAACCAAGAC	120
DB	61	CGGCTCGTCTGCTGTTGTTGTCAGTAATGATTTCTGAGACCGGTTACAAACCAAGAC	120
QY	121	CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG	180
DB	121	CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG	180
QY	181	CCATCATCGTACCGAATGTTGGAGGATCCCCAGAGACCTCCAGTGTACCCCGAGCGG	240
DB	181	CCATCATCGTACCGAATGTTGGAGGATCCCCAGAGACCTCCAGTGTACCCCGAGCGG	240

QY	241	CAGAGTGAAGGACGGGTATATGAAGGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC	300
DB	241	CAGAGTGAAGGACGGGTATATGAAGGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC	300
QY	301	ACCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCCCTGCCTCTGGGTCTTTAAGCAC	360
DB	301	ACCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCCCTGCCTCTGGGTCTTTAAGCAC	360
QY	361	AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCCATGCC	420
DB	361	AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCCATGCC	420
QY	421	ATCTTGAACCTGACAGAGACCCAGGAGAGATACTTCCATATTCAGAGGAGCAAGC	480
DB	421	ATCTTGAACCTGACAGAGACCCAGGAGAGATACTTCCATATTCAGAGGAGCAAGC	480
QY	481	GCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG	540
DB	481	GCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG	540
QY	541	AGACCTTACTTTAGGAAGATGGAACACAGSATGCTGCTGCTCATCTCCGAGGGTGT	600
DB	541	AGACCTTACTTTAGGAAGATGGAACACAGSATGCTGCTGCTCATCTCCGAGGGTGT	600
QY	601	CCGAGGCCACTGTTGAGTGGTGCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA	660
DB	601	CCGAGGCCACTGTTGAGTGGTGCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA	660
QY	661	GGCCCTGCTGTTGTCAGAAAGGAGAAAGTACTTTCATGAGTTGTTGGAGACAGATC	720
DB	661	GGCCCTGCTGTTGTCAGAAAGGAGAAAGTACTTTCATGAGTTGTTGGAGACAGATC	720
QY	721	AGATGCTGCTAGAAATGCACCTGGCGCGAATGCACCAAGCTGTTCACCATAGATCTA	780
DB	721	AGATGCTGCTAGAAATGCACCTGGCGCGAATGCACCAAGCTGTTCACCATAGATCTA	780
QY	781	AACAGGCTCTCAGAGCACACTGCCCCAGTTTCTGAAAGTGGGGGAACCCCTGTGG	840
DB	781	AACAGGCTCTCAGAGCACACTGCCCCAGTTTCTGAAAGTGGGGGAACCCCTGTGG	840
QY	841	ATCAGGTGAAGGCCATCCATGTGAACCATGGATTTCGGGCTCAGCTGGGAGTGAAGAC	900
DB	841	ATCAGGTGAAGGCCATCCATGTGAACCATGGATTTCGGGCTCAGCTGGGAGTGAAGAC	900
QY	901	AAAGCCCTGGAGGAGGCGCAGTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	960
DB	901	AAAGCCCTGGAGGAGGCGCAGTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	960
QY	961	ATTCGGATTCTCTTGGCCCTTTGTCTTCCTGGGAAGAACGACACCGGATATTACACC	1020
DB	961	ATTCGGATTCTCTTGGCCCTTTGTCTTCCTGGGAAGAACGACACCGGATATTACACC	1020
QY	1021	TGCTCTTCTCAAGCACCCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1080
DB	1021	TGCTCTTCTCAAGCACCCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1080
QY	1081	ATAACGCTACAGCTCGCAAGAGAGATGAAATTCACCCGTCACGAAAGTCTTCGCTTC	1140
DB	1081	ATAACGCTACAGCTCGCAAGAGAGATGAAATTCACCCGTCACGAAAGTCTTCGCTTC	1140
QY	1141	TCAGTCAAGTAAAGCGTACCCACGAATCCGATGACGAGTGTCTTCTCAAGACCTCA	1200
DB	1141	TCAGTCAAGTAAAGCGTACCCACGAATCCGATGACGAGTGTCTTCTCAAGACCTCA	1200
QY	1201	TTTCTTGTGAACAGAGAGGCTGGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
DB	1201	TTTCTTGTGAACAGAGAGGCTGGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
QY	1261	AAGAAACAGCCAGAGAGTACATATCTATGCAGAAAATGATGACGCCCTTCACCAAA	1320
DB	1261	AAGAAACAGCCAGAGAGTACATATCTATGCAGAAAATGATGACGCCCTTCACCAAA	1320
QY	1321	ATGTTACGCTGAATATAGAAAACCTCAAGTGTAGCAATGCTCTCAGCCAGCCAG	1380

1321 ATGTTACAGCTGAATATAGAAAGAAACCTCAAGTGTAGCAATTCCTCAGCCACCAG 1380
1381 GGGTCTGTTCTCTGATGGCTACCGCTACCGCTCTGAGCTGGAAGAGTGTTCGGAC 1440
1381 GGGTCTGTTCTCTGATGGCTACCGCTACCGCTCTGAGCTGGAAGAGTGTTCGGAC 1440
1441 AAATCTCCCAATTGACGAGGAAATCCAGAGGAGTTGGAAATAAAAGGCTACACA 1500
1441 AAATCTCCCAATTGACGAGGAAATCCAGAGGAGTTGGAAATAAAAGGCTACACA 1500
1501 AAAGTGTGTTGGCAGTGGGTGCGAGCAGTACTCTAAATATAGTGAGGCGGGAAAGG 1560
1501 AAAGTGTGTTGGCAGTGGGTGCGAGCAGTACTCTAAATATAGTGAGGCGGGAAAGG 1560
1561 CTTCTGGTCAATGCTGCGGTACAAATTCATGCGGACGCTCTGCGAAACCATCTTTTA 1620
1561 CTTCTGGTCAATGCTGCGGTACAAATTCATGCGGACGCTCTGCGAAACCATCTTTTA 1620
1621 AACTCACCAGGCCCTTCCCTTTCATCCAAAGACACATCTCTCTATGCGACCATGGG 1680
1621 AACTCACCAGGCCCTTCCCTTTCATCCAAAGACACATCTCTCTATGCGACCATGGG 1680
1681 CTCGTCTCCCTTCATTGTTGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCA 1740
1681 CTCGTCTCCCTTCATTGTTGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCA 1740
1741 TTTAGTACGAGTACGAGTCCAGATGATCCAGGTGACTGGGCCCTCGTATACGAGTAC 1800
1741 TTTAGTACGAGTACGAGTCCAGATGATCCAGGTGACTGGGCCCTCGTATACGAGTAC 1800
1801 TTTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860
1801 TTTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860
1861 TTAGAGTTTGGAGGCTCTGGGCTGCGGCTTTTCGGAGGAGTGAAGACGCCACGGCC 1920
1861 TTAGAGTTTGGAGGCTCTGGGCTGCGGCTTTTCGGAGGAGTGAAGACGCCACGGCC 1920
1921 TATGGCATTTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAAA 1980
1921 TATGGCATTTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAAA 1980
1981 GCTGACAGCTGTGAAAGAAAGCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGCTGTGAAAGAAAGCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACAAATCGTAATCTGCTGGGGATGCACACTGTGAGGGCAGTGTACTTG 2100
2041 CACCATGACAAATCGTAATCTGCTGGGGATGCACACTGTGAGGGCAGTGTACTTG 2100
2101 ATTTTGAATATTTGTTGCTATGGTACCTCCTCACTACTAAGAGTAAAGAGAGAAG 2160
2101 ATTTTGAATATTTGTTGCTATGGTACCTCCTCACTACTAAGAGTAAAGAGAGAAG 2160
2161 TTTACAGACATGAGACAGATTTTAAAGGACATAATTTTCAGTTCTTACCCTACTTTC 2220
2161 TTTACAGACATGAGACAGATTTTAAAGGACATAATTTTCAGTTCTTACCCTACTTTC 2220
2221 CAGGACATTTCAATTCACAGATGCTGGTTACAGAGAGTTTCAGTTACACCCGCTTG 2280
2221 CAGGACATTTCAATTCACAGATGCTGGTTACAGAGAGTTTCAGTTACACCCGCTTG 2280
2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATATGAA 2340
2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATATGAA 2340
2341 AACCAAGAGGCTGCGAGAAGAGGAGGAGGAGATTGAACTGCTGAGCTTTGAGAC 2400
2341 AACCAAGAGGCTGCGAGAAGAGGAGGAGGAGATTGAACTGCTGAGCTTTGAGAC 2400
2401 CTTCTTTGCTTTCGTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460

2401 CTCTTTGTTTTCGTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTGT 2460
2461 GTCCACAGAGACCTGCGACCCAGGAATGTTGGTCAACCCACGGGAAGTGGTCAAGATC 2520
2461 GTCCACAGAGACCTGCGACCCAGGAATGTTGGTCAACCCACGGGAAGTGGTCAAGATC 2520
2521 TGTGACTTTGGACTGCGCCGAGACATCCTCAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580
2521 TGTGACTTTGGACTGCGCCGAGACATCCTCAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580
2581 GCACGGCTGCCGTTGAAGTGGACCGACGAGCTTTATTGAAGGATCTACACAACT 2640
2581 GCACGGCTGCCGTTGAAGTGGACCGACGAGCTTTATTGAAGGATCTACACAACT 2640
2641 AAGAGTACGTCGTCGTCCTACGGCATCTCTCTGGAGATATTTTACCTGGGTGTAAC 2700
2641 AAGAGTACGTCGTCGTCCTACGGCATCTCTCTGGAGATATTTTACCTGGGTGTAAC 2700
2701 CTTTACCCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2760
2701 CTTTACCCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2760
2761 ATGAGCAGCCATTTCTATGCCAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820
2761 ATGAGCAGCCATTTCTATGCCAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820
2821 TTTGACTCAAGGAAGGGCCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880
2821 TTTGACTCAAGGAAGGGCCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880
2881 GCAGGCGCAGAAAGCATGTATCAGAAATCATCTCATCTACCAACAGCGCGCCCT 2940
2881 GCAGGCGCAGAAAGCATGTATCAGAAATCATCTCATCTACCAACAGCGCGCCCT 2940
2941 CAGCAGAGAGCGGGCTCAGAGCCAGTGCACAGCGCCAGTGCACAGCGCCAGAGTTCACAGAGAA 3000
2941 CAGCAGAGAGCGGGCTCAGAGCCAGTGCACAGCGCCAGTGCACAGCGCCAGAGTTCACAGAGAA 3000
3001 AGAAGTTAGCGAGGAGCGCTTGGACCCCGCCACCTAGCAGCTGAGCCGCGAGAGCA 3060
3001 AGAAGTTAGCGAGGAGCGCTTGGACCCCGCCACCTAGCAGCTGAGCCGCGAGAGCA 3060
3061 AGATTAGCTGCGCTCTGAGGAAGCGCCCTACAGCGCTTCTGCTGCGTGGACTTTCTCT 3120
3061 AGATTAGCTGCGCTCTGAGGAAGCGCCCTACAGCGCTTCTGCTGCGTGGACTTTCTCT 3120
3121 AGATGCTGCTGCCATTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180
3121 AGATGCTGCTGCCATTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180
3181 GCGGGAGAGCAATTAATGAGACTTGTGGTGAGCCCGCTACCCCTGGGGGCTTTTCCACG 3240
3181 GCGGGAGAGCAATTAATGAGACTTGTGGTGAGCCCGCTACCCCTGGGGGCTTTTCCACG 3240
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATCTCTGTAATACGTGAACAA 3300
3241 AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATCTCTGTAATACGTGAACAA 3300
3301 ACCAAACCCCTTTTTCGTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360
3301 ACCAAACCCCTTTTTCGTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360
3361 AATGACTATGTAACCTTTTTCATCTATTTAGTGAATATTTTATGATGGAATAAATCTTTC 3420
3361 AATGACTATGTAACCTTTTTCATCTATTTAGTGAATATTTTATGATGGAATAAATCTTTC 3420
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 3

US-09-919-408-3

1088 CTACGAGTCGCAAGAGATGATGAATTTGACCGGTACGAAAGTTCTCTCTCAGTCA 1147
1112 CTACCAATTCAGTCAAGATTAAGAAATGACCAATATGAAGAGTTTCTCTCTGTC 1171
1148 GGTTTAAAGCGTACCCAGATTCGATGACGATGATCTCTCTCAAGCCTCATTTCCCT 1207
1172 GGTTTAAAGCGTACCCAGATTCAGATGATGATGATGATGATGATGATGATGATG 1231
1208 GTGACAGAGAGCGTGGAGGATGGTACAGATATCTAAATTTCCGATCATAGAACA 1267
1232 GTGACCAAAAGGCTTGTATACGATACAGATATCCAAAGTTTGGCAATCAAGACCC 1291
1268 AGCCAGGAGATATATCTATGACAGAAATGATGACGCCAGTTTCCACCAAAATGTTCA 1327
1292 AGCCAGGAGATATATCTATGACAGAAATGATGATGATGATGATGATGATGATGATG 1351
1328 CGCTGAATATAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCGGCTCCT 1387
1352 CGCTGAATATAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCGGCTCCT 1411
1388 GTTCTCTGATGCTTACCGCTTACCCCTTGTGACCTGGAAGATGTTCCGACAAATCTC 1447
1412 GTTCTCGATGATACCCATACCCTTGTGACCTGGAAGATGTTCCGACAAATCTC 1471
1448 CCAATTCAGGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGAAAAGTGT 1507
1472 CCAACTGCACAGAGATACAGAGGAGTCTGGAAATAGAAAGGCTAACAGAAAAGTGT 1531
1508 TTGGCCAGTGGTGTGCGAGCTACTCTAAATATAGTGTAGGCGCGGAAAGGCTCTGG 1567
1532 TTGGACAGTGGTGTGCGAGCTACTCTAAACATGAGTGAAGCCATAAAAGGTTCTGG 1591
1568 TCAATGCTGTGCGTACAAATCTATGGGACGCTTGTGGAACCATCTTTTAAACTCAC 1627
1592 TCAAGTGTGTACATAAATTCCTTGGCAGCTTGTGAGAGGATCTTTTAAACTCTC 1651
1628 CAGGCGCTTCCCTTTCATCCAAAGACATCTCTCTATCGGACCATTTGGGCTCTGTC 1687
1652 CAGGCGCTTCCCTTTCATCCAAAGACATCTCTCTATCGAACAATTTGGTGTGTC 1711
1688 TCCCTTCATTTGTTCTCATTTGTTGATCGCCACAAATACAAAAGCAATTTAGGT 1747
1712 TCCCTTCATTTGTTCTCATTTGTTGATCGCCACAAATACAAAAGCAATTTAGGT 1771
1748 ACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1807
1772 ATGAAAGCCAGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1831
1808 TTGACTTCAGGACTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAACTTAGAGT 1867
1832 TTGATTCAGAGAAATGAATATGATCTCAAAATGGAGTTTCCAAAGAAATTTAGAGT 1891
1868 TTGGAAGCTCTCGGGCTCGGCTTTCGGGAGGTGATGAACGCCAGCGCTATGGCA 1927
1892 TTGGAAGCTACTAGATCAGTGTGTTTGGAAAGTGTGAACCGACAGCTTATGGAA 1951
1928 TTGTAAGACCGGAGTCTCAATTCAGTGTGCGGTGAAGATGCTTAAAGAAAGCTGACA 1987
1952 TTAGCAAAACAGGAGTCTCAATTCAGTGTGCGGTCAAAATGCTGAAAGAAAAGCAGCA 2011
1988 GCTGTGAAAAGAACTCTCATGTCGGAGCTCAAAATGATGACCCAGCTGGACACCATG 2047
2012 GCTCTGAAAGAGGAGCTCATGTCAGAACTCAAGATGATGATGATGATGATGATGATG 2071
2048 ACAACATCGTGAATCTGTGGGGGATGACACTGTGAGGGCAGTGTACTGATTTTG 2107
2072 AGAATATTGTGAACCTGTGGGGGCGTGACACTGTGAGGACCAATTTACTTGATTTTG 2131
2108 AATATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2167
2132 AATACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2191

2168 GGACATGGACAGAGATTTTAAAGAAATAAATTTCACTTCTTCTTCTTCTTCTTCTTCTT 2227
2192 GGACTTGGACAGAGATTTTCAAGAAACACAAATTTCACTTCTTCTTCTTCTTCTTCTTCT 2251
2228 ATTCAAATTCAGATGCTCTGTTTACGAGAAATTTCACTTCTTCTTCTTCTTCTTCTTCTTCT 2287
2252 ATCCAAATTCAGATGCTCTGTTTCAAGAGAAATTTCACTTCTTCTTCTTCTTCTTCTTCTTCT 2311
2288 TCTCAGGTTTCAATGGAATTTCAATTTCTCAAGATGATGATGATGATGATGATGATGATGAT 2347
2312 TCTCAGGTTTCAATGGAATTTCAATTTCTCAAGATGATGATGATGATGATGATGATGATGAT 2371
2348 AGAGGCTGTCAGAAAGAGGAGGAGATTTGAACGTTGCTGAGCTTTTGAAGACTTCTTTT 2407
2372 AAAGGCTG-----GAAGAGAGGAGGACTTGAATGCTTACATTTGAAGATCTTCTT 2425
2408 GCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTGAGATTTCAAGTCTGCTGTGTCACA 2467
2426 GCTTTGCGATATCAAGTTGCCAAAGAAATTTCTGGAATTTAAGTCTGTGTTTCA 2485
2468 GAGACTGTCAGCAGGAATGTTGTCACCCACCGGAAGTGGTGAAGATCTGTGACT 2527
2486 GAGACTGTCGCGCAGGAAGTGTGTCACCCACCGGAAGTGGTGAAGATCTGTGACT 2545
2528 TTGGACTGTCGCGCAGGACATCTGAGGACTTCCAGCTACGTCGTCAGGGGCAACGACGCG 2587
2546 TTGGATTTGCTCGAGATATCATGATGATTTCCAACTATGTTGTCAGGGGCAATGCCGCTC 2605
2588 TCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATCAAGATG 2647
2606 TGCCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2665
2648 AGCTGCTGCTTACGCGATCTTCTCTGCGAGATATTTTCACTGGGTGTGAACCTTACC 2707
2666 ATGCTGCTCATATGAATATTTACTGTGGAAATTTCTTCACTGGTGTGAATCTTACC 2725
2708 CTGGCATTCCTGTCGAGCTTAATCTTATAAATGATTCAGAGTGGATTTAAATGAGC 2767
2726 CTGGCATTCCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2785
2768 AGCCATTTCTATGCCAGAGGAGATATCTTCTTGAATGCAATCTCTGCTGGCTTTTCACT 2827
2786 AGCCATTTTATGCTACAGAAATATATACATTAATGCAATCTCTGCTGGCTTTTCACT 2845
2828 CAAGGAGGCGCATCTTCCCAACCTGATCTATTTTAGGATGTCAGCTGGCAGG 2887
2846 CAAGGAGGCGCATCTTCCCTTAATTTGACTTCGTTTGGATGTCAGCTGGCAGATG 2905
2888 CAGAAGAAG-----CATGTATCAGAACATCCATCATCTACC 2924
2906 CAGAAGAAGGATGATCAGAAATGTCGATGGCGGTGTTTCCGAATGTCCTCACACCTTACC 2965
2925 AAAACAGCGCGCGCTTCCAGCAGAGAGCGCGCTCAGAGC--CCAGTCCCCACAGCGCCAGG 2983
2966 AAAACAGCGCGCTTCCAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 3025
2984 T--GAGATTCACAGAAAGAACTTACGAGGAGGCTTGGACCCCGC-----ACCTAG 3038
3026 TCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACTATCCCTAA 3085
3039 CAGGCTGTAGACCGCAGAGCCAGATAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3096
3086 CAGGCTGTAGATTTACCAAAACAAAGATTAATTTTCACTTAAAGAAATATATATATCAAC 3145
3097 CGTGCTTCTGCTGAGCTTCTCTAGATGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3149
3146 TCTGCTTCCAGCAGCTTCTCTAGAGCGCTCTGGTGTACTCTTCTTCTTCTTCTTCTTCTTCT 3205
3150 CTCTTATATAAATCAAACTCTCTCTCAGAGCGGGAGGCCAATATAGACTTGTG 3209
3206 CTTTGTAAATCAAAATCATCTCTGTCACAGGCGGAGGAGCTGATTAATGAACCTTTATTG 3265
3210 TGAGCCCGCTACCTTGGGGCGCTTTCACAGGACTTTCAGGGGAAAGCCATGTATCTGAAA 3269

3266 GAGCATTGATCTGCATCCAGGCGCTTCTCAGGCGGCTTGAGTGAATGTGTAACCTGAAG 3325
3270 TATAGTATATCTGTGAATACGTGAACAAACACCCGCTTTTGTGCTAAGGGAAGC 3329
3326 TACAGTATATCTGTGAATACATAAAACAAA-----AGCATTTTGTGAAGGAAGC 3378
3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
3379 TAATATGATTTT--AAGTCTATGTTTAAATACTATGTAACCTTTTTCAGCTATTAG 3436
3390 TGATATATTTATGATGGAATAAATCTTCTACTGTAAATAAATAAATAAATAAATAA 3449
3437 TGATATATTTATGCGTGGGAATAAATTTCTACTACAGAAATAAATAAATAAATAA 3496
3450 AAAA 3453
3497 AAAA 3500

RESULT 4
US-09-872-136-3
Sequence 3, Application US/09872136
Patent No. US20020119545A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..138
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-872-136-3
Query Match 64.5%; Score 2227.2; DB 10; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 8 GCCTACCGCGGCTCCGGAGGCGCATGGCGCGCTTGGCGAGCGACGCGCGGCTGC 67
DB 35 GGGGACCCCGGCTCCGGAGGCGCATGGCGCGCTTGGC---GGCGGACGCGGACCGCTGC 91
QY 68 TGCTGCTTGTGTTTGTGCTAGTAAATGATTTGTAGACCTTTACAAACCAAGACCTGCCGTG 127
DB 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCGTG 151
QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAAATGGCTCATCAGCGGGAAGCCATCAT 187
DB 152 TGATCAAGTGTGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 211
QY 188 CGTACCGAATGCTGGAGGATCCCGAAGACCTCCAGTGTACCCGAGCGCGGACAGTG 247
DB 212 CATATCCCATGTTATCAGAAATCCCGAAGACCTCGGCTGCGTTGAGACCCAGAGCT 271
QY 248 AAGGACGCTATATGAGCGGCGACCGTGGAGGTGCCGAGCTGTGGTTCATCACCCTGC 307
DB 272 CAGGACAGTGTACGAAAGCTGCCGCTGTGGAAGTGTATCTGTCTTCCATCACACTGC 331
QY 308 AAGTGCAGCTCGCCACCCCGAGGACCTTTCCTGCTCTGGGTCTTTAAGCAGAGCTCCC 367
DB 332 AAGTGCCTGCTGATGCCCGGAGCAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 368 TGGGCTGCCAGCGCGCACTTTGATTTACAAACAGAGGAATCTGTTTCCATGCCCATCTTGA 427
DB 392 TGAATTCGCCAGCCACATTTGATTTACAAACAGAGAGAGTGTGTTCCATGTCATTTTGA 451
QY 428 AGCTGACAGAGACCCAGGAGGAGATACCTACTTCCATATTCAGAGCGAAGCGGCCAACT 487
DB 452 AAATGACAGAAACCCAAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511
QY 488 ACACAGTACTGTTTCAGAGTGAATGTAGAGATACAGAGCTGTATGTCTTAAAGAGACTT 547
DB 512 ACACAAATATGTTTACAGTGAATAAAGAAATACCTGCTTTACATTAAGAGAGACTT 571
QY 548 ACTTTAGGAGATGGAAACCCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
DB 572 ACTTTAGAAATAAGAAACCCAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631
QY 608 CCACTGTGGAGTGGTGTCTGCGAGCTCCACAGGGAAGCTGTAAAGAGAAAGAGCCCTG 667
DB 632 CGATCGTGAATGGGTGCTTTGGGATTCACAGGGGGAAGCTGTAAAGAGAAAGTCCAG 691
QY 668 CTGTTGTACAGAAAGGAGGAAAGGTACTTTCATGATTTGTTTCGGAACAGACATCATGCT 727

Db 692 CTGTTGTTAAAAAGGAGGAAAAAGTCTTCATGAATTAATTTGGACGGACATAAGGTGCT 751
QY 728 GTGCTAGAAATGCATGGCCCGGAATGCACCAAGCTGTTCACCATAGATCTAAACACAGG 787
Db 752 GTGCCAGAAATGAATGGGCGAGGAATGCACAGGCTGTTCACAAATAGATCTAAATCAAA 811
QY 788 CTCCTCAGACACACTGCCCCAGTTTATTCCTGAAAGTGGGGAACCCCTTGTGATCAGT 847
Db 812 CTCCTCAGACCACATGCCACAATTAATTCCTTAAGTAGGGGAACCCCTTATGTAAGGT 871
QY 848 GTAAGGCCATCCATGTGAACCATGGATTCCGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907
Db 872 GCAAAGCTGTTCAATGTGAACCATGGATTCCGGCTCACCTGGGAATTTAGAAAAACAAAGCAC 931
QY 908 TGGAGGAGGGAGCTACITTTAGATGAGTACCTACTCCACAACAGGACCATGATTCGGA 967
Db 932 TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTTCAACAACAGAACTATGATACGGA 991
QY 968 TTCTCTTGCCCTTTGTCTTCCGTGGGAAGAACACACCCGATATTACACCTGCTCTT 1027
Db 992 TTCTGTGTCTTTGTATCATCAGTGGCAAGAAACACACCCGATACTACACTGTGTCCT 1051
QY 1028 COTCAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAAAGGGTTTATAAAGC 1087
Db 1052 CTTCAAAGCATCCAGTCAATCAGCTTGGTTACCATCGTAGAAGGGATTTATAAATG 1111
QY 1088 CTACAGCTCGCAAGAAGATGATAAATGACCCGTACGAAAGTTCTCTCTCAGTCA 1147
Db 1112 CTACCAATTTCAAGTAGAATGATAAATGACCAATATGAAGAGTTTCTTTCTCTGTCA 1171
QY 1148 GGTTTAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCAAGCTCATTTCCCTT 1207
Db 1172 GGTTTAAAGCCTTACCACCAATACAGATGACGTGGACCTTCTCTGAAATCATTTCCCTT 1231
QY 1208 GTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTCGATCATATAAGAAC 1267
Db 1232 GTGAGCAAAAGGCTTGATACGGATACAGCATATCCAAATTTTCCATCATAGCACCC 1291
QY 1268 AGCCAGGAGAGTACATATCTATGCGAAAAATGATGACCCCGATTCACCAAAATGTTCA 1327
Db 1292 AGCCAGGAGAAATATATTTCCATGACGAAAAATGATGCCAAATTTACCAAAATGTTCA 1351
QY 1328 CCGTGAATATAAGAAAGAACTCAAGTCTAGCAAAATGCCTCAGCAGCCAGCGCTCT 1387
Db 1352 CCGTGAATATAAGAAAGAACTCAAGTCTCGCAAGCATCGGCAAGTCAGGCGTCT 1411
QY 1388 GTTCTCTGATGGCTACCGCTACCGTCTTGGACCTGGAAGAGTGTTCGGACAAATCTC 1447
Db 1412 GTTCTCGGATGATACCATTAACCATCTTGGACCTGGAAGAGTGTTCAGACAAATCTC 1471
QY 1448 CCAATTGCAGGAGGAAATCCCAAGAGGATTTGGAATAAAAGGCTAACAGAAAAAGTGT 1507
Db 1472 CCAACTGCACAGAGAGATCAGCAAGGAGTCTGGAATAGAAGGCTAACAGAAAAAGTGT 1531
QY 1508 TTGGCCAGTGGTTCGACACTACTCTAAATATGATGAGCGCGGGAAGGCTTCTCG 1567
Db 1532 TTGGACAGTGGGTGTCGACAGTACTCTAAACATGATGAAGGCTTAAAGGGTCTCTCG 1591
QY 1568 TCAAAATGCTGTGCTACAAATTTCTATGGCAGCTTCTTGGCAAAACCATCTTTTAAACTCAC 1627
Db 1592 TCAAGTGTGCTACATATCCCTTGGCACATCTTGTGAGAGATCCCTTTTAAACTCTC 1651
QY 1628 CAGGCCCTTCCCTTTTATCCCAAGACAACTCTCTTCTATGCGACCAATGGGCTCTGTC 1687
Db 1652 CAGGCCCTTCCCTTTTATCCCAAGACAACTCTCTTCTATGCAACAAATTTGGTGTGTC 1711
QY 1688 TCCCTTCAATGTTGTTCTCATTTGTTGATCTGCGCAAAAATACAAAAAGCAATTTAGGT 1747
Db 1712 TCCCTTCAATGTTGTTTAAACCTGCTAATTTGTCAAGTACAAAAAGCAATTTAGGT 1771
QY 1748 ACGAGAGTCAAGTGCAGATGATCCAGTGCACGTGGCCCTGGATACAGTACTTCTAGC 1807
Db 1771 TCCCTTCAATGTTGTTTAAACCTGCTAATTTGTCAAGTACAAAAAGCAATTTAGGT 1807

Db 1772 ATGAAAGCCAGCTACAGATGGTACAGGTGACCGGCTCCTCAGATAATGAGTACTTCTACG 1831
QY 1808 TTACACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTTCCCAGAGAGACTTAGAGT 1867
Db 1832 TTGATTTTCAGAGAAATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAAAATTTAGAGT 1891
QY 1868 TTGGGAAGGTCCCTGGGCTCTGGGCTTTCGGAGGAGTGATGAACCCACACCGCTTATGCA 1927
Db 1892 TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCAACACCTTATGGA 1951
QY 1928 TTAGTAAAAGCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAAGAGAAAGCTGACA 1987
Db 1952 TTAGCAAAAACAGGAGTCTCAATCCAGTTGCCGTCAAAATGCTGAAAAGAAAAAGCAGACA 2011
QY 1988 GCTGTGAAAAGAGAGCTCTCATGTCGAGCTCAAAATGATCACCACCTGGGACACCATG 2047
Db 2012 GCTCTGAAAAGAGAGGACCTCATGTCAGAACTCAAGATGATGACCCAGCTGGGAGGCCAG 2071
QY 2048 ACAACATCGTGAATCTGCTGGGGCATGCACACTGTTCAGGGCCAGTGTACTGATTTTGTG 2107
Db 2072 AGAATATTGGAACCTGCTGGGGCGTGCACACTGTTCAGGACCAATTTACTTGATTTTG 2131
QY 2108 AATATTGTTGCTATGTTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGAAAGTTTCA 2167
Db 2132 AATACTGTTGCTATGTTGATCTTCAACTATCTAAGAAGTAAAGAGAAAAATTTCA 2191
QY 2168 GGCATGAGCAGAGATTTTAAAGAACATTAATTCAGTTCTTACCTTCTTCCAGGCAC 2227
Db 2192 GGCATTTGGACAGAGATTTTCAAGGAACACAAATTTTACCTTCTTCCACATCTTCCA 2251
QY 2228 ATTCAAATTCACAGCTGCTGTTTCCAGAGAAGTTCAGTTACACCCCGCTTGGATCAGC 2287
Db 2252 ATCCAAATTCACAGCTGCTGTTTCAAGAGAAGTTCAGATACACCCCGCTTGGATCAA 2311
QY 2288 TCTCAGGGTTCAATGGGAATTCATTCATTCGAAGATGAGATGAATATGAAACACAGA 2347
Db 2312 TCTCAGGGCTTCATGGGAATTCATTTTCACTCTGAAGATGAATTTGAATATGAAACCAA 2371
QY 2348 AGAGCTGGCAGAGAAGAGAGAGAGAGATTTGAACGTGCTCAGCTTTGAAGACCTCTT 2407
Db 2372 AAAGGCTG-----GAAGAAGAGGAGACTTGAATGTCTTACATTTGAAGATCTCTT 2425
QY 2408 GCTTTGGTACCAGTGGCCAAAGGCATGGAATTTCTCGAGTTCAAGTGTGTGTCCACA 2467
Db 2426 GCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTGTGTGTACA 2485
QY 2468 GAGACCTGGCAGCAGGAATGTGTGTCACCCACGGGAAGGTGGTGAAGATCTGTGACT 2527
Db 2486 GAGACCTGGCCGCGCAGGAACGTGCTTGTACCCACGGGAAGTGGTGAAGATATGTGACT 2545
QY 2528 TTGACCTGGCCCGCAGACATCCTGAGCAGCTCCAGCTAGTCTCAGGGGCAACGACGCG 2587
Db 2546 TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTGAGGGCAATGCCGCTC 2605
QY 2588 TGGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAACTCAAGAGTG 2647
Db 2606 TGCTGTAAATGATGGCCCCCGAAAGCCTGTTTGAAGGATCTACACCATTAAGAGTG 2665
QY 2648 ACCTCTGGTCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
Db 2666 ATGCTGTGTCATATGGAATATTAAGTGGGAATCTTCTCAGCTTGGTGAATCCTTACC 2725
QY 2708 CTGCAATCTCTGTCGAGGCTTAACCTCTATAAAGTATGATGATGATTTTAAATGGAGC 2767
Db 2726 CTGGCATTCGGTGTGATGCTAACTTCTACAACTGATTCAAAATGGATTTTAAATGGATC 2785
QY 2768 AGCAATCTATGCCACAGAGAGGATATCTTGTGAATGCAATCTGCTGGCTTTTGGTACT 2827
Db 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCTGCTGGCTTTTGGTACT 2845
QY 2828 CAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTATAGGATGTCAGCTGGCAGAGG 2887
Db 2846 CAAGGAACCGGCATCTTCCCTAATTTGACTTCTGTTTATAGGATGTCAGCTGGCAGAG 2905

2888 CAGAGAAG-----CATGTATCAGAACATCCATCATACC 2924
|||||
2906 CAGAGAAGCATGTATCAGAAATGGGCGGTTCGGAATGCTCTCACACTACC 2965
|||||
2925 AAAACAGGGCGCCCTCAGCAGAGGGCGGCTCAGAGC-CCAGTGCACACAGCCGAGG 2983
|||||
2966 AAAACAGGGCGACCTTTCAGCAGAGAGATGGATTGGGGGTACTCTCTCCGCGGCTCAGG 3025
|||||
2984 T-GAGATTCACAGAGAAGAGTTAGCAGAGAGGCGCTGGACCGGCC-ACCTAG 3038
|||||
3026 TCGAAGATTCTGAGAGGACAAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
|||||
3039 CAGGCTGTAGACGGCAGAGCCAGATTAAGCTTCGCTCT--GAGGAAGCGCCCTACAGCG 3096
|||||
3086 CAGGCTGTAGATTACCAACAGAGTAATTTTCATCACTAAAGAAATCTATTATCAAC 3145
|||||
3097 CGTCTGCTCGGCGGACTTTCTCTAGATGCTGTGCGCATTAATC-----CAAAGTGA 3149
|||||
3146 TCGTCTTCCACGAGACTTTTCTCTAGAACCGCTCTCGTTTACTCTTTGTTTCAAAGGGA 3205
|||||
3150 CTTCTATAAATCAAACTCTCTCGCACAGGGGGGAGAGCCAAATAGAGACTTGTGG 3209
|||||
3206 CTTTGTAAATCAAACTCATCTGTCTACAAAGGAGGAGGCTGATAATGAACCTTTATTG 3265
|||||
3210 TGAGCCCGCTACCTCGGGGCTTTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269
|||||
3266 GAGCATGTATCTGCAATCCAGGCTTCTCAGGCGGCTTGAGTGAATTTGTACCTGGAAG 3325
|||||
3270 TATAGTATATTTCTGTAATAGTGAACAAACAAACCCGTTTTTTTGTCTAAGGGAAGC 3329
|||||
3326 TACAGTATATTTCTGTAATATACATAAAACAA-----AGCATTTTGTCTAAGGGAAGC 3378
|||||
3330 TAAATATGATTTTAAATATATGTTTAAATATGATGTAACCTTTTTCATCTATTAG 3389
|||||
3379 TAATATGATTTT--AAGTCTATGTTTAAATATATGTAATTTTTCAGCTATTAG 3436
|||||
3390 TCATATATTTTATGATGAAATAAATCTTCTACTGTAATAAATAAATAAATAAATAA 3449
|||||
3437 TGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAAATAAATAAATAA 3496
|||||
3450 AAAA 3453
|||||
3497 AAAA 3500
|||||

RESULT 5

US-09-982-610-22
; Sequence 22, Application US/09982610
; Patent No. US20020146420A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; Bennett, Brian D.
; Goeddel, David
; Lee, James M.
; Matthews, William
; Tsai, Siao Ping
; Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982.610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: F0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-9881
TELEFAX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3120 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-982-610-22

Query Match 60.3%; Score 2081.8; DB 10; Length 3120;
Best Local Similarity 80.9%; Pred. No. 0;
Matches 2513; Conservative 0; Mismatches 557; Indels 37; Gaps 6;
QY 48 GCACGACGACGCGCGCTGCTGCTGTGTTTTCAGTATGATTTCTTGAGACCGT 107
Db 15 GCACGACGCGCGCGCTGCTGCTGTGTTTTCAGTATGATTTCTTGAGACCGT 74
QY 108 TACAAACACAGACCTGCTGCTGTGATCAAGTGTGTTTAAATCAGTATGATTTCTTGAGACCGT 167
Db 75 TACAAATCAAGATCTGCTGCTGTGATCAAGTGTGTTTAAATCAGTATGATTTCTTGAGACCGT 134
QY 168 ATCAGCGGGAACCATCATCTGATCCCAATGGTGCAGAGATCCCAAGAACCTCCAGTG 227
Db 135 ATCAGTGGGAATCATCATCATATCCATGGTATCAGATCCCGGAGACCTCGGCTG 194
QY 228 TACCCGACGCGCGCAGAGTGAAGGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGA 287
Db 195 TCGCTTGGAGACCGCGCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATG 254
QY 288 GTCTGGTCCATCACCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTCTCTGCTCTG 347
Db 255 ATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCCAGGGAACATTTCTCTGCTCTG 314
QY 348 GGTCTTTAAGCACAGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAAACAGAGGAAT 407
Db 315 GGTCTTTAAGCACAGCTCCCTGATTTGCCAGCCACATTTTGATTTACAAAACAGAGGAGT 374
QY 408 CGTTTCCATGGCCATCTTGAACGTGACAGAGACCCAGGAGGAAATACCTTACATAT 467
Db 375 TGTTCATGCTCATTTTGAATAATGACAGAAACCCAAAGCTGGAGAAATACCTTTTAT 434
QY 468 TCAGAGGAGCGCGCCCACTACACAGTACTGTTTCACAGTGNATGTAGAGATACACAGCT 527
Db 435 TCAGAGTGAAGCTACCAATTTACCAATATTTTACAGTGTAGTATAGAAATACCTCTCT 494
QY 528 GTATGTCTAAGGAGACCTTACTTTAGGAAGATGAAAACCCAGGATGCCTGCTGCTGAT 587
Db 495 TTACACATTAGAAGACCTTACTTTAGAAATGAGAAACCCAGGAGGAGGAGGAGGAG 554
QY 588 CTCCGAGGCTTCCCGAGCGCCACTGTGGAGTGGGTGCTCTGACAGTCCCAAGGAGGAG 647
Db 555 ATCTGAGAGCGTCCAGAGCGGATCCCTGGAATGGTGTCTTGGCAGTTCACAGGGGAGG 614
QY 648 CTGTAAGAAGAGGCGCTCTGTGTCAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 707
Db 615 CTGTAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 674

Qy 708 CGGAACAGACATCAGATGCTGTGTAGAAATGCACCTGGCGCGGAATGCACCAAGCTGTT 767
Db 675 TGGCAGCGACATAGGTGCTGTGCCAGAAATGAATGGCGAGGAAATGCCACGCGCTGTT 734
Qy 768 CACCATAGATCTAAACAGAGCTCTCAGAGACACACTGCCCGAGTTATTCCTGAAAGTGGG 827
Db 735 CACAATAGATCTAAATCAAACTCTCAGACACACATTCGCCAATATTTCTTTAAAGTAGG 794
Qy 828 GGAACCTTGTGGATAGGTCTAGGCCATCCATGCTGAACCATGATTCGGGCTCACCTG 887
Db 795 GGAACCTTATGGATAGGTCTAGAGCTGTTCATGTGAACCATGATTCGGGCTCACCTG 854
Qy 888 GGAGCTGGAACACAAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACTACTCCAC 947
Db 855 GGAATTAGAAACAAAGCCTCGAGGAGGCGAATCTTTGAGATGAGTACTACTTCAAC 914
Qy 948 AAACAGGACCATGATTCGGATTCCTTGGCCCTTTGTGTCTTCCTGGGAAAGAGACGAC 1007
Db 915 AAACAGAACTATGATACGGATTCCTTGTCTTTGTATCATCAGTGGCAAGAAAGACAC 974
Qy 1008 CGGATATTACACCTGCTCTCTCAAGACCCAGCCAGTCAGGTTGGTGACCATCCT 1067
Db 975 CGGATCTACACTGTTCTCTTCAAGCATCCCGAGTCAATCAGCTTTGGTTACCATCGT 1034
Qy 1068 AGAAAAAGGGTTTATAACCGCTACAGCTCGCAAGAGAGATGATAAATTGACCCCTACGA 1127
Db 1035 AGAAAGGGATTTATAATGCTACCAATTCAGTTCAAGTATGAATGAGCAATATGA 1094
Qy 1128 AAAGTCTGCTTCAGTCAGGTTTAAAGCTTACCAGCAATCCGATCGATCGATGATCTT 1187
Db 1095 AGAGTTTGTGTTCTGTCAGGTTTAAAGCTTACCACAAATCAGATGATGAGTGGACCTT 1154
Qy 1188 CTCTCAGGCTCATTTCTTGTGAACAGAGAGGCTTGAGGATGGTACACATATCTAA 1247
Db 1155 CTCTCGAAATCATTTCTTGTGACAAAGGGTCTTGATTAACGGATACAGCATATCCAA 1214
Qy 1248 ATTTTGGCATCATAGAACAGCCAGGAGATACATATTTCTATCGAGAAAATGATGAGCC 1307
Db 1215 GTTTTGAATCATAGAACAGCCAGGAGATATATATTTCCATCGAGAAATGATGATGC 1274
Qy 1308 CCAGTTCCACAAATGTTACGCTGGAATATAGAAAGAAACCTCAAGTGTAGCAATGTC 1367
Db 1275 CCAATTTACCAAAATGTTACGCTGTATATAGAAAGAAACCTCAAGTCTCGCAAGAC 1334
Qy 1368 CTCAGCCAGCAGCGCTCTCTCTGTGAGTACCCGCTACCCCTTTGGACCTGAA 1427
Db 1335 TTGGGCAAGTCAGCGCTCTTTCTGGATGGATACCATTTACCATCTTGGACCTGAA 1394
Qy 1428 GAAGTGTTCGACAAATCTCCCAATTCACGAGGAGAAATCCAGAGAGTTTGGAAATA 1487
Db 1395 GAAGTGTTCAGACAAATCTCCCAATTCACAGAGAGATACAGAGAGTCTGGAATAG 1454
Qy 1488 AAAGGCTTAACAGAAAGTGTTCGACAGTGTGCGTACAAATTCATGCGCAGCTTTGGA 1547
Db 1455 AAAGGCTTAACAGAAAGTGTTCGACAGTGTGCGTACAAATTCATGCGCAGCTTTGGA 1514
Qy 1548 GGCGGGAAGGGCTTCTGTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1607
Db 1515 AGCCATAAAAAGGGTTCCTGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1574
Qy 1608 AACCATCTTTTAACTACACGAGCCCTTCCCTTTTCAATCAAGACACATCTCTCTCTA 1667
Db 1575 GAGATCTTTTAACTCTCAGCGCCCTTCCCTTTTCAATCAAGACACATCTCTCTCTA 1634
Qy 1668 TGGACCATTTGGGCTCTGCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCGCAAA 1727
Db 1635 TGCAACAAATTTGGTGTCTCTCTCTCATTTGCTGTTTAACTGCTTAAATTTGTCACAA 1694
Qy 1728 ATACAAAGCAATTTAGGTACGAGATGTCAGTGTGAGTATGATGATGATGATGATGATG 1787
Db 1695 GTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTACAGGTGACCGGATCTC 1754
Qy 1788 GGATACGAGTACTTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTT 1847

Db 1755 AGATTATGAGTACTTCTACGTTGATTTTCAGAAATATGAATATGATGTCAAAATGGAGTT 1814
Qy 1848 CCCGAGAGAACTTTAGAGTTTGGGAAAGTCTCTGGGCTCTGGGCTTTCCGGGAGGTGAT 1907
Db 1815 TCCAAGAGAAATTTAGAGTTTGGGAAAGTCTAGTACAGTGTCTTTTGGAAAGTATGAT 1874
Qy 1908 GAACCCACGGCTATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGAT 1967
Db 1875 GAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGTTACCGTCAAAAT 1934
Qy 1968 GCTAAAAGAGAAAGTGTGACAGCTGTGAAAAGAAAGCTCTATCTCGAGCTCAAAATGAT 2027
Db 1935 GCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGACTCAAGATGAT 1994
Qy 2028 GACCCACCTGGGACACACATGACAACTGCTGTAATCTGCTGGGGGCAATGACACTGTGAGG 2087
Db 1995 GACCCAGCTGGGAAGCCACAGAAATTTGTAACCTGCTGGGGGCTGACACTGTGAGG 2054
Qy 2088 GCCAGTGTACTTGTATTTTGAATATTTGCTATGCTGACCTCTCAACTACCTAAGAAG 2147
Db 2055 ACCAATTTACTGATTTTGAATTTGCTATGCTGATCTCTCACTATCTAAGAAG 2114
Qy 2148 TAAAGAGAGAAAGTTTCACAGGACATGGACAGAGATTTTAAAGAAACATAAATTCAGTTC 2207
Db 2115 TAAAGAGAGAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCACTT 2174
Qy 2208 TTACCTTACTTTCCAGGACATTTCAATTTCCAGCATGCTGTTCCAGGAGAACTTCAGTT 2267
Db 2175 TTACCCCACTTTTCCAATCACATCCAAATTTCCAGCATGCTGTTCAAGAGAACTTCAGAT 2234
Qy 2268 ACACCCGCTTGGATCAGCTCTCAGGTTTCAATGGAAATTCAAATTCATCTCAAGATGA 2327
Db 2235 ACACCCGACTCGGATCAATCTCAGGCTTCAATGGAAATTCATTTCACTCTCAAGATGA 2294
Qy 2328 GATTGAATATGAAACACCAAGAGGCTGCGACAGAGAGGAGAGAGATTTGAACCTGCT 2387
Db 2295 AATTCAATATGAAACCAACCAAAAGGCTG-----GAAGAGAGGAGGACTTGAATGCTG 2348
Qy 2388 GACGTTTGAAGACCTCTTTGCTTTGCTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGA 2447
Db 2349 TACATTTGAAGATCTTCTTTGCTTTGCTATCAAGTTGCCAAAGAAATGGAATTTCTGGA 2408
Qy 2448 GTTCAAGTGTGCTGCTCCAGAGAGCTGCGACGAGAGATGCTGCTGCTACCCACCGGAA 2507
Db 2409 ATTTAAGTGTGCTGCTCCAGAGAGCTGCGCCGCGAGAGAGCTGCTGCTACCCACCGGAA 2468
Qy 2508 GGTGTGGAAGATCTGTGACTTTGGACTGGCCGCGAGACATCTCTGAGGAGCTTCCAGCTACGT 2567
Db 2469 AGTGTGGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTCACACTATGT 2528
Qy 2568 GCTCAGGCGCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2627
Db 2529 TGTCAAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2588
Qy 2628 GATCTACACAACTCAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2687
Db 2589 CATCTACACCAATTAAGATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2648
Qy 2688 ACTGGTGTGAACCTTTACCTTGGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2747
Db 2649 ACTTGTGTGATCTTACCTTGGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2708
Qy 2748 GAGTGGATTTAAATGGAGAGCCATTTCTATGCCAGAGAGGATATCTTTTGTATGCA 2807
Db 2709 AAATGGATTTAAATGGATCAGCCATTTTATGTACAGAGAAATATACATTTAATGCA 2768
Qy 2808 ATCTGCTGGGCTTTTGTACTCAAGNAGGCGCCATCTTCCCCCAACTGACTTCATTTT 2867
Db 2769 ATCTGCTGGGCTTTTGTACTCAAGNAGGCGCCATCTTCCCCCAACTGACTTCATTTT 2828
Qy 2868 AGGATGTGAGTGTGGAGGCGCAGAGAG- - - - -CATGTATC 2904

;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7538
;; LENGTH: 645
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-7538

Query Match 13.5%; Score 466.6; DB 9; Length 645;
Best Local Similarity 84.5%; Pred. No. 8.7e-118;
Matches 550; Conservative 0; Mismatches 94; Indels 7; Gaps 2;
QY 2035 CTGGACACCATGCAACATCTGCTGGGGGCGATGCACACTGTCTAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAATATTGTGNACCTGCTGGGGGCGTGCACACTGTCTAGGACCAATT 60
QY 2095 TACTTGATTTTGAATATTGTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTGCTATGCTGATCTTCTCAACTATCTAAGAGTAAAGA 120
QY 2155 GAGAAGTTTCACAGACATGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCTT 2214
Db 121 GAAAAATTTTCAAGSACTTGGACAGAGATTTTCAAGGAACACAATTTTCACTTCTTACCCC 180
QY 2215 ACTTTCCAGGCACATTTCAAAATTCAGCATGCTGCTTCCAGGAGAGTTTCAGTTACACCG 2274
Db 181 ACTTTCACATCATCTCAAGGCTTTCAGGAAATTCATTTCACTCTGGAAGTGAATTCGA 240
QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTGAA 2334
Db 301 TATGAAACCAACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
QY 2395 GAAGACCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2454
Db 355 GAAGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
QY 2455 TCAGTGTCTCCACAGACCTGCGCAGCCAGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
Db 415 TCAGTGTCTCCACAGACCTGCGCAGCCAGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
QY 2515 AAGATCTGTGACTTTGGACTGGCCGCGAGACATCTCTGAGCGACTCAGCTAGCTGCTGAGG 2574
Db 475 AAGATCTGTGACTTTGGACTGGCCGCGAGACATCTCTGAGCGACTCAGCTAGCTGCTGAGG 534
QY 2575 GGCACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2633
Db 535 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
QY 2634 CACAATCAAGAGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2684
Db 595 CACCATTAAGAGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645

RESULT 8
US-05-796-692-8085
;; Sequence 8085, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; FILE REFERENCE: 2077,001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8085
;; LENGTH: 610
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-8085

Query Match 13.2%; Score 456.4; DB 9; Length 610;
Best Local Similarity 85.1%; Pred. No. 5.3e-115;
Matches 524; Conservative 0; Mismatches 86; Indels 6; Gaps 1;
QY 2035 CTGGACACCATGCAACATCTGCTGGGGGCGATGCACACTGTCTAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGGCGTGCACACTGTCTAGGACCAATT 60
QY 2095 TACTTGATTTTGAATATTGTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTGCTATGCTGATCTTCTCAACTATCTAAGAGTAAAGA 120
QY 2155 GAGAAGTTTCACAGACATGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCTT 2214
Db 121 GAAAAATTTTCAAGSACTTGGACAGAGATTTTCAAGGAACACAATTTTCACTTCTTACCCC 180
QY 2215 ACTTTCCAGGCACATTTCAAAATTCAGCATGCTGCTTCCAGGAGAGTTTCAGTTACACCG 2274
Db 181 ACTTTCACATCATCTCAAGGCTTTCAGGAAATTCATTTCTGAAGATGAGATTGAA 2334
QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTGAA 2394
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTCAAGATGAAATTGAA 300
QY 2395 TATGAAACCAACCAAGAGCTGCGCAGAGAGAGGAGAGATTTGACGCTGCTGACGCTTT 2394
Db 301 TATGAAACCAACCAAGAGCTGCGCAGAGAGAGGAGAGATTTGACGCTGCTGACGCTTT 354
QY 2395 GAAGACCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2454
Db 355 GAAGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
QY 2455 TCGTGTCTCCACAGACCTGCGCAGCCAGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
Db 415 TCGTGTCTCCACAGACCTGCGCAGCCAGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474


```
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8927

Query Match      13.28; Score 454.6; DB 9; Length 605;
Best Local Similarity 85.3%; Pred. No. 1.7e-114;
Matches 521; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGACATCGTGAATCTGTGGGGCGTGCACACTGTGAGGCCAGTG 2094
Db 1 CTGGGAGCCACGAGATATTTGTAACCTGTCTGGGGCGTGCACACTGTGAGGCCAAT 60

QY 2095 TACTTGATTTTGAATATTTGCTATGCTGACCTTCTCAACTACCTAAGAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGCTATGCTGACCTTCTCAACTATCTAAGAAGTAAAGA 120

QY 2155 GAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCACTTACCC 2214
Db 121 GAAAATTTTCACAGGACATGGACAGAGATTTTCAAGGACACAAATTTCACTTTTACCCC 180

QY 2215 ACTTTCCAGGACATTTCAATTTCCAGCATGCTGTTCACGAGAAGTTTCACTTACCCG 2274
Db 181 ACTTTCCAAATCATCTCCAAATTTCCAGCATGCTGTTCACGAGAAGTTTCACTTACCCG 240

QY 2275 CCCTTGGATGACCTGTGAGTGGGATTTTAAAGGAACATAATTTCACTTGAAGATGAA 2334
Db 241 GACTCGGATCAATCTCAGGGCTTTCATGGGATTTCACTTGAAGATGAAATGAA 300

QY 2335 TATGAAAACCAAGAGAGGCTGCGCAAGAAAGAGGAGGATTTTGAACGTGCTGAGCTTT 2394
Db 301 TATGAAAACCAAGAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTT 354

QY 2395 GAAGACCTTCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTTCTTTGCTTTGCTATATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTAAG 414

QY 2455 TCCTGTTGCCACAGACCTGGACCCAGGAGATTTGTTGGTCCACCCACGGGAGGTGGTG 2514
Db 415 TCCTGTTGCCACAGACCTGGCCGACGAGACCTGCTTGTACCCACGGGAAAGTGGTG 474

QY 2515 AAGATCTGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTAGCTCGTCAGG 2574
Db 475 AAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTGTCAGG 534

QY 2575 GGCAACCCAGCGTGGCGGTGAAGTGGATGGCACCAGGAGCTTATTTGAAGGATCTAC 2634
Db 535 GGCAATGCCGCTGTGCTGTGAATGATGGCCCGCCGAAAGCCTGTTTGAAGGATCTAC 594

QY 2635 ACAATCAAGAG 2645
Db 595 ACCATTAAGAG 605
```

RESULT 11

```
US-10-040-862-8927
; Sequence 8927, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
```

```
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8927
```

```
Query Match      13.28; Score 454.6; DB 9; Length 605;
Best Local Similarity 85.3%; Pred. No. 1.7e-114;
Matches 521; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGACATCGTGAATCTGTGGGGCGTGCACACTGTGAGGCCAGTG 2094
Db 1 CTGGGAGCCACGAGATATTTGTAACCTGTCTGGGGCGTGCACACTGTGAGGCCAAT 60

QY 2095 TACTTGATTTTGAATATTTGCTATGCTGACCTTCTCAACTACCTAAGAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGCTATGCTGACCTTCTCAACTATCTAAGAAGTAAAGA 120

QY 2155 GAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCACTTACCC 2214
Db 121 GAAAATTTTCACAGGACATGGACAGAGATTTTCAAGGACACAAATTTCACTTTTACCCC 180

QY 2215 ACTTTCCAGGACATTTCAATTTCCAGCATGCTGTTCACGAGAAGTTTCACTTACCCG 2274
Db 181 ACTTTCCAAATCATCTCCAAATTTCCAGCATGCTGTTCACGAGAAGTTTCACTTACCCG 240

QY 2275 CCCTTGGATGACCTGTGAGTGGGATTTTAAAGGAACATAATTTCACTTGAAGATGAA 2334
Db 241 GACTCGGATCAATCTCAGGGCTTTCATGGGATTTCACTTGAAGATGAAATGAA 300

QY 2335 TATGAAAACCAAGAGAGGCTGCGCAAGAAAGAGGAGGATTTTGAACGTGCTGAGCTTT 2394
Db 301 TATGAAAACCAAGAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTT 354

QY 2395 GAAGACCTTCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTTCTTTGCTTTGCTATATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTAAG 414

QY 2455 TCCTGTTGCCACAGACCTGGACCCAGGAGATTTGTTGGTCCACCCACGGGAGGTGGTG 2514
Db 415 TCCTGTTGCCACAGACCTGGCCGACGAGACCTGCTTGTACCCACGGGAAAGTGGTG 474

QY 2515 AAGATCTGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTAGCTCGTCAGG 2574
Db 475 AAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTGTCAGG 534

QY 2575 GGCAACCCAGCGTGGCGGTGAAGTGGATGGCACCAGGAGCTTATTTGAAGGATCTAC 2634
Db 535 GGCAATGCCGCTGTGCTGTGAATGATGGCCCGCCGAAAGCCTGTTTGAAGGATCTAC 594

QY 2635 ACAATCAAGAG 2645
Db 595 ACCATTAAGAG 605
```



```
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7976
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
us-10-040-862-7976

Query Match      13.0%; Score 448.8; DB 9; Length 597;
Best Local Similarity 85.4%; Pred. No. 6.5e-113; Indels 6; Gaps 1;
Matches 514; Conservative 0; Mismatches 82;

QY 2035 CTGGGACACCATGACATCGTGAATCTGTGGGGCATGCACACTGTCTAGGCGCCAGTG 2094
Db 1 CTGGGAAGCCAGGAGATATGTGAACCTGTGGGGCGTGCACACTGTCTAGGACCAATT 60

QY 2095 TACTTGATTTTGAATATTTGTTGATGGTGACCTCTCAACTACCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGTTGATGGTGATCTCTCAACTATCTAGAAAGTAAAGA 120

QY 2155 GAGAAGTTTCACAGGACATGCAGAGATTTTAAAGAACATATTTTCAGTCTTACCT 2214
Db 121 GAAAAATTTTCAGGACATTTGTAACCTGTCTGGGGGGTGCACACTGTCTAGGACCAATT 180

QY 2215 ACTTTCCAGGCACATTCACATCTGCTGATGACCTCTCAAGAGAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTCAAGAGAGTTCAGATACACCG 240

QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAAGATGAGATTGAA 2334
Db 241 GACTCGGATCAATCTCTAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

QY 2335 TATGAACACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGTGCTGACGTTT 2394
Db 301 TATGAACACCAAGAGAGGCTG-----GAAGAGAGAGAGAGGCTTGAATCTGCTTACATTT 354

QY 2395 GAAGACCTCTCTTGTCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTCTTGTCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAAATTGAA 414

QY 2455 TCCTGTGTCACAGAGACCTGGCAGCAGGAGATGTTGTCACCCACGGGAAGGTGGTG 2514
Db 415 TCCTGTGTTTCACAGAGACCTGGCCGCGCAGGAACTGCTTGTACCCACGGGAAGGTGGTG 474

QY 2515 AAGATCTGACATTTGACATGGCCCGGAGACATCTCTGAGGACCTCCAGCTGCTCAGG 2574
Db 475 AAGATATGTCAGTTGGATTGGCTCGAGATATCATGATGATTCCTCAACTATGTTGTCAGG 534

QY 2575 GGCAAGCAGCGCTGCGGTTGAAGTGGATGGCAGCAGGAGCTTATTTGAAGGATCTAC 2634
Db 535 GGCAATGCCGCTGCGCTGTAAATGGATGGCCCGGAGAGGCTGTTTGAAGGATCTAC 594

QY 2635 AC 2636
Db 595 AC 596

RESULT 14
us-09-796-692-8273
; Sequence 8273, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
```

```
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n-A,T,C or G
us-09-796-692-8273

Query Match      13.0%; Score 448; DB 9; Length 601;
Best Local Similarity 84.8%; Pred. No. 1.1e-112;
Matches 515; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGCACACATCGTGAATCTGTGGGGCATGCACACTGTCTAGGCGCCAGTG 2094
Db 1 CTGGGAAGCCAGGAGATATTTGTAACCTGTCTGGGGGGTGCACACTGTCTAGGACCAATT 60

QY 2095 TACTTGATTTTGAATATTTGTTGATGGTGACCTCTCAACTACCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGTTGATGGTGATCTCTCAACTATCTAGAAAGTAAAGA 120

QY 2155 GAGAAGTTTCACAGGACATGGACAGAGATTTTAAAGAACATPAATTTCACTGCTTACCT 2214
Db 121 GAAAAATTTTCAGGACATTTGTAACAGAGATTTTCAAGGAACACAAATTTTCAGTTTTCACCC 180

QY 2215 ACTTTCCAGGCACATTCAAATTCAGCATGCTGTTTCCAGGAGAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTCAGATACACCG 240

QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAAGATGAGATTGAA 2334
Db 241 GACTCGGATCAATCTCTAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

QY 2335 TATGAACACCAAGAGAGGCTGGCAGAGAGAGAGAGATTTTGAACGTGCTGACGTTT 2394
Db 301 TATGAACACCAAGAGAGGCTG-----GAAGAGAGAGAGGCTTGAATCTGCTTACATTT 354

QY 2395 GAAGACCTCTCTTGTCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTCTTGTCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAAATTGAA 414

QY 2455 TCCTGTGTCACAGAGACCTGGCAGCAGGAGATGTTGTCACCCACGGGAAGGTGGTG 2514
Db 415 TCCTGTGTTTCACAGAGACCTGGCCGCGCAGGAACTGCTTGTACCCACGGGAAGGTGGTG 474

QY 2515 AAGATCTGACATTTGACATGGCCCGGAGACATCTCTGAGGACCTCCAGCTGCTCAGG 2574
Db 475 AAGATATGTCAGTTGGATTGGCTCGAGATATCATGATGATTCCTCAACTATGTTGTCAGG 534
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:00:43 ; Search time 4534.48 Seconds
(without alignments)
12332.831 Million cell updates/sec

Title: US-09-919-408-1
Perfect score: 3453
Sequence: 1 GCGGCTGGCTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_fod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	655.8	19.0	765	9	AI323253 mp88b02.y
2	556.8	16.1	618	9	AA120050 mp88b02.r
3	531.8	15.4	604	12	BF523018 UI-R-C2p-
4	502.4	14.5	747	9	AI323643 mp88b02.x
5	486.6	14.1	522	10	BB644407 BB644407
6	452.6	13.1	621	13	BI360262 BI360262

C	7	445	12..9	445	10	BE651447	BE651447 UI-M-BH3-
	8	432.8	12.5	714	13	BI461248	BI461248 603206574
	9	426.2	12.3	515	14	BQ556177	BQ556177 H4038808-
C	10	423	12.3	484	13	BM117950	BM117950 L0858803-
	11	354.4	10.3	507	13	BM484050	BM484050 537756 MA
	12	340.8	9.9	460	10	AV713950	AV713950 AV713950
	13	333.6	9.7	3312	11	AK004947	AK004947 Mus muscu
	14	324.8	9.4	5174	11	BC026713	BC026713 Mus muscu
	15	314	9.1	394	12	BG145808	BG145808 un98g07.y
C	16	294	8.5	457	9	AA996414	AA996414 UI-R-C0-h
	17	277.6	8.0	357	9	AI556371	AI556371 UI-R-C2p-
C	18	273	7.9	424	9	AI180146	AI180146 EST223884
	19	266	7.7	502	12	BF557871	BF557871 UI-R-C0-h
	20	265.8	7.7	455	13	BM147250	BM147250 TCAAP1010
	21	254	7.4	558	10	BE679272	BE679272 dg61c06.y
C	22	248.2	7.2	375	9	AA851570	AA851570 EST194338
	23	244.8	7.1	660	10	BE047675	BE047675 tz40G09.y
C	24	227.8	6.6	386	9	AI178876	AI178876 EST222558
	25	227.6	6.6	407	10	AV596131	AV596131 AV596131
	26	226.6	6.6	739	12	BF159205	BF159205 601766656
	27	226.4	6.6	294	10	BB247268	BB247268 BB247268
	28	224.8	6.5	666	9	AL047744	AL047744 DKFZP586P
	29	224.6	6.5	503	13	BI481351	BI481351 H2RPE-165
	30	219.4	6.4	666	13	BG934591	BG934591 SK1-0903
	31	219.2	6.3	286	10	BB332140	BB332140 BB332140
	32	216	6.3	848	13	BI107537	BI107537 602893883
	33	215	6.2	940	9	AI325368	AI325368 mi30h09.y
	34	214.4	6.2	702	12	BF163671	BF163671 601769775
C	35	213.4	6.2	469	12	BF189757	BF189757 235453 MA
	36	212	6.1	696	10	BE260791	BE260791 601150285
	37	212	6.1	794	13	BI111563	BI111563 602895917
	38	211	6.1	832	13	BI694328	BI694328 602347787
	39	209.8	6.1	531	12	BF320739	BF320739 u255h04.y
	40	207.6	6.0	575	12	BF930052	BF930052 MR2-NT013
	41	205	5.9	696	14	BM951127	BM951127 UI-M-EH0-
	42	202.8	5.9	665	13	BG922846	BG922846 602823693
	43	201.4	5.8	412	12	BE853746	BE853746 ux20412.y
	44	201	5.8	290	10	BB282068	BB282068 BB282068
	45	199.4	5.8	669	13	BJ012204	BJ012204 BJ012204

ALIGNMENTS

RESULT 1
AI323253
LOCUS
DEFINITION
AI323253
mp88b02.y1 Soares.thymus.2NBMT Mus musculus cDNA clone IMAGE:576267
5' similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN); PDGF
gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF
(MOUSE);, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

AI323253 765 bp mRNA linear EST 23-DEC-1998
mp88b02.y1 Soares.thymus.2NBMT Mus musculus cDNA clone IMAGE:576267
5' similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN); PDGF
gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF
(MOUSE);, mRNA sequence.
AI323253 GI:4057682
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 765)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314.286 1810
Email: mouseest@watson.wustl.edu

FEATURES

BASE COUNT
ORIGIN

Best Local Similarity 94.0%; Pred. NO. 1.1E+03;
Matches 699; Conservative 0; Mismatches 36; Indels 4; Gaps 2;

Db 1 GGGCTTCTGTCAAATGCTGTGCGTACAAATTCATGGGCACGCTTCGGAACCATCTTT 60

Db 61 TTAAACTCACCAGGCCCTTCCTTTCATCCAGACACATCTCCTTCTATGCGACCAT 120

Db 121 GGGCTGTGTCFCCCCCTCATTTGTTGTTCTTCATTGTTGATCTGCCACAAATACAAAAAG 180

[illegible]

D0
D0
241 IACIIICACGIIIGACIICAGGGGACATGGTAIGGCCIIAAGIUGGAGIICCCTCGAGGAGS 309

QV 1858 AACTTACACTTTGGGAGGTCCTCTGGGGGTCTGGGGTTCGGCAGGCTGATCAACGGCCAG 1917

1918 GCCTATGGCATTAGTAAACGGGAGTCTCAATTCAAGGTGGCGGTGAAGATGCTATAAAGAG 1977

QY 1978 AAAGCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCCACTG 2037

2038 GGACACCATGACAACATCGTGAATCTGCTGGGGGCATGCACACTGTCAGGGCCAGTGTAC 2039

Qy 2098 TTGATTTTGAATATTGTGCTATGGTGACCTCCTCAACTACCTAAGAAAGTAAAAGAGAG 2157

RESULT 2
AA120050

ACCESSION
VERSION

TITLE
JOURNAL

FEATURES

PAGE COUNT

BASE COUNT
ORIGIN

Query Match 16.1%; Score 556.8; DB 9; Length 618;
 Best Local Similarity 97.7%; Pred. No. 1.6e-86;
 Matches 586; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 1558 GGGCTTCGGTCAAAATGCTGCGTACAAATCTATGGCAGCTCTTGCGAAACCAATCTTT 1617
 Db 1 GGGCTTCGGTCAAAATGCTGCGTACAAATCTATGGCAGCTCTTGCGAAACCAATCTTT 60

QY 1618 TTAACCTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTATGCGACCAT 1677
 Db 61 TTAACCTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTATGCGACCAT 120

QY 1678 GGGCTCTCTCCCTTCATCTGTTCTCATTTGTTGATCTGCCACAAATACAAAAG 1737
 Db 121 GGGCTCTCTCCCTTCATCTGTTCTCATTTGTTGATCTGCCACAAATACAAAAG 180

QY 1738 CAATTTAGGTACGAGATCAGTGCAGATGATCCAGTGGCCCTCGATTAACGAG 1797
 Db 181 CAATTTAGGTACGAGATCAGTGCAGATGATCCAGTGGCCCTCGATTAACGAG 240

QY 1798 TACTTCTACGTTGACTTCAGGAGCTATCAATATGACCTTAAGTGGAGTTCGCGAGAG 1857
 Db 241 TACTTCTACGTTGACTTCAGGAGCTATCAATATGACCTTAAGTGGAGTTCGCGAGAG 300

QY 1858 AACTTAGAGTTTGGGAAGTCTCGGGTCTGGCGCTTTTCGGGAGGTTGATGAACGCCACG 1917
 Db 301 AACTTAGAGTTTGGGAAGTCTCGGGTCTGGCGCTTTTCGGGAGGTTGATGAACGCCACG 360

QY 1918 GCCTATGGCATTAATAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAG 1977
 Db 361 GCCTATGGCATTAATAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAG 420

QY 1978 AAAGCTCAGAGCTGTAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTG 2037
 Db 421 AAAGCTCAGAGCTGTAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTG 480

QY 2038 GGACACCATGACACATCGTAATCTGCTGGGGCATGCACACTGTACAGGCCAGTGTAC 2097
 Db 481 GGACACCATGACACATCGTAATCTGCTGGGGCATGCACACTGTACAGGCCAGTGTAC 540

QY 2098 TTGATTTTGAATATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2157
 Db 541 CTGATTTTGAATATGTTG-ATGGTGACCT-CTCTACTTCCGTAGAAGTTAAGAGAGAG 598

RESULT 3
 BF523018
 LOCUS
 DEFINITION UI-R-C2p-rg-f-10-0-UI.r1 UI-R-C2p Rattus norvegicus cDNA clone
 ACCESSION BF523018
 VERSION BF523018.1 GI:11631033
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 604)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at
 LLNL (infoimage.llnl.gov). IMAGE ID- 1793119
 Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
 1..604
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-rg-f-10-0-UI"
 /clone_lib="UI-R-C2p"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 166 a 132 c 167 g 139 t
 ORIGIN

Query Match 15.4%; Score 531.8; DB 12; Length 604;
 Best Local Similarity 93.7%; Pred. No. 3.3e-82;
 Matches 554; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1991 GTGAAAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACCATGACA 2050
 Db 8 GTGAAAAGAGAGCGCTCATGCTGAGCTCAAAATGATGACCCACCTGGGACCATGACA 67

QY 2051 ACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGACTTGTATTTTGAAT 2110
 Db 68 ACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGACTTGTATTTTGAAT 127

QY 2111 ATTGTTGCTATGGTGACCTCTCAACTACCTAAGAAGTAAAGAGAGAGTTTCACAGGA 2170
 Db 128 ATTGTTGCCATGCTGACCTGCTCAACTACCTAAGAAGTAAAGAGAGAGTTTCACAGGA 187

QY 2171 CATGACAGAGATTTTAAAGAACATATTTTACCTTCTTACCTTCTTCCAGGCACAT 2230
 Db 188 CGTGACAGAGATTTTAAAGAACATATTTTACCTTCTTACCTTCTTCCAGGCACAT 247

QY 2231 CAAATTCACAGATGCTGTTCCAGAGAGTTTACCTTACACCCGCCCTTGGATCAGCTCT 2290
 Db 248 CAACTCCAGTATGCCGGGTTCCAGAGAGTTTACAGATATACCCGCCCTTGGATCAGCTCT 307

QY 2291 CAGGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATATGAAAACCAAGA 2350
 Db 308 CAGGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATATGAAAACCAAGA 367

QY 2351 GGCTGGCAGAGAGAGAGGAGGAGATTTGAACGCTGCTGACGTTTGAAGACCTCTTGTCT 2410
 Db 368 GGCTGGAAGAGAGAGGAGGAGGAGATTTGAACGCTGCTGACGTTTGAAGACCTCTTGTCT 427

QY 2411 TTGGGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTGTTGTCCACAGAG 2470
 Db 428 TTGGGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTGTTGTCCACAGAG 487

[illegible]

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-2 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
waqi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001).
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

[illegible]

```

Location/Qualifiers
1. .522
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B230315G04"
/clone_lib="RIKEN full-length enriched, adult male corpora
quadrigemina"
/sex="male"
/tissue_type="corpora quadrigemina"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCCAAGAGCTCTTTTCTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTAAATTAATATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."
116 a 149 c 151 g 106 t

```

	Query Match	14.1%	Score 486.6;	DB 10;	Length 522;
	Best Local Similarity	99.0%;	Pred. No. 2.3e-74;		
	Matches 500;	Conservative 0;	Mismatches 4;	Indels 1;	Gaps 1;
QY	1	GCGCCCTGGCTACCGCGCGCTCCGGAGGCCATCGCGCGCTTGGCGCAGCGACGACCGG	60		
Db	18	GCGCCCTGGCTACCGCGCGCTCCGGAGGCCATCGCGCGCTTGGCGCAGCGACGACCGG	77		
QY	61	CGGCTGCTGCTGCTTGTGTTTTGTCAGTAATGATCTTGAGAGCCGTTACAAACCAGAC	120		

Db	78	CGGCTGCTGCTGCTGTGTGTTTTGTCAGTAATGATTTCTGAGACCGGTTACAAACCAAGAC	137
Qy	121	CTGCCTCTGATCAAGTGTGTTTTTAATCAGTCATCAGAACAAATGGCTCATCAGCGGGAAG	180
Db	138	CTGCCTCTGATCAAGTGTGTTTTTAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAG	197
Qy	181	CCATCATCTGACCGAATGGTGCAGGAGTCCCGAAGAACCTCCAGTGTACCCCGAGSGCG	240
Db	198	CCATCATCTGACCGAATGGTGCAGGAGTCCCGAAGAACCTCCAGTGTGCCCGAGSGCG	257
Qy	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300
Db	258	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	317
Qy	301	ACCTCGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCGCCTCTGGGTCTTTTAAGCAC	360
Db	318	ACCTCGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCGCCTCTGGGTCTTTTAAGCAC	377
Qy	361	AGTCCCTGGGTGCCAGCGGCACATTGTATTTACAAACAGAGGAATCGTTTCCATGGCC	420
Db	378	AGTTCCTGGGTGCCAGCGGCACATTGTATTTACAAACAGAGGAATCGTTTCCATGGCC	437
Qy	421	ATCTTGAACTGCACAGAGACCCAGGCGAGGAGATACCTACTCCATATTCAGAGGAACGC	480
Db	438	ATCTTGAACTGCACAGAGACCCAGGCGAGGAGATACCTACTCCATATTCAGAGGAACGC	497
Qy	481	GCCAACTACACAGTACTGTTTCACAG	505
Db	498	GCC - ACTACACAGTACTGTTTCACAG	521

[illegible]

REFERENCE	TITLE	JOURNAL	COMMENT
1 (pages 1 to 621) Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W. and Keele, J.W.	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine	Unpublished (2000)	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGCG
 Plate: 139 row: H column: 2
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 1. .621

```

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pcwv SPORN6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```

```
BASE COUNT      190 a      132 c      141 g      158 t
ORIGIN
Query Match      13.1%; Score 452.6; DB 13; Length 621;
Best Local Similarity 83.2%; Pred. No. 1.6e-68;
Matches 515; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1702 GTTCTCATTTGCTTGCATCGGCACAAATACAAAAGCAATTTAGTACGAGAGTGCAGCTG 1761
DB 3 GTTTTAAACATGCTAAATTTGTACAAAGTACAAAAGCGGATTCGGTACGAAAGCCAGCTG 62
QY 1762 CAGATGATCCAGGTGAGTGGCCCTGGATACAGAGTACTTCTACGCTTCAGGAG 1821
DB 63 CAGATGGTCCAGGTGAGCGGGTCCCTGGATACAGTACTTCTACGCTTCAGAGAA 122
QY 1822 TATGAATATGACCTTAAGTGGGAGTTCGCCGAGAGAACTTAGAGTTTGGGAAGTCTCG 1881
DB 123 TATGAATATGACCTCAAAATGGGAGTTTCCCGAGGAAATTTAGAATTTGGGAAGTCTCG 182
QY 1882 GGGTCTGGCGTTTTCGGGAGGTGATGACGCCAGCGCTATGGCATAGTAAACGGGA 1941
DB 183 GGATCGCGCGCTTTTGGAAAGTGTGAACGCACTGCTATGGCAATCAGTAAACTGGA 242
QY 1942 GTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGCAGAGCTGTGAAAAGAA 2001
DB 243 GTGTCAATCCAGGTGGCAGTCAAAATGCTGAAGAAAGGCGAGAGAGCTCGGAACGAGAG 302
QY 2002 GCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGACACCATGACAACTGCGTGAAT 2061
DB 303 GCTCTCATGCTGTAATCAAAATGATGACCCACCTGGGACGAGGAGAAATAGTGAAC 362
QY 2062 CTGCTGGGGGATGACACTGTCAGGGCCAGTGTACTTGATTTTGAATATTGTGCTAT 2121
DB 363 CTGCTGGGGGCTGACCCCTGTGAGGGCCCATTTACTTGATTTTGAATATTGTGCTAC 422
QY 2122 GGTGACCTCTCAACTACCTAAGAAGTAAAGAGAGAACTTTCACAGGACATGGACAGAG 2181
DB 423 GGTGATCTCTCACTATCTAGAGTAAAGAGAGAAATTTCTATAGGACATGGACCGAG 482
QY 2182 ATTTTAAAGAACATAATTTCACTTCTTACCTTCTTCCAGGACATTCAAATTCACG 2241
DB 483 ATTTTCAAGGAACATAATTTCACTTCTTCTTCTTCTTCCAAATCACACCCAAATTCAGT 542
QY 2242 ATGCTGTTTCAGGAGAACTTCACTACACCCGCTTGGATCAGCTCTCAGGTTCAAT 2301
DB 543 ATGCGGGGTTCAAGAGAACTTCAATACACCCAGACTCGGATCTTCTCAGGATTCAT 602
QY 2302 GGAATTCATTCATCTG 2320
DB 603 GGAATTCATTCATCTG 621
RESULT 7
BE651447/c      445 bp      mRNA      linear      EST 06-SEP-2000
LOCUS      UI-M-BH3-atk-d-06-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
DEFINITION      UI-M-BH3-atk-d-06-0-UI 5', mRNA sequence.
ACCESSION      BE651447
VERSION      BE651447
KEYWORDS      EST.
SOURCE      house mouse..
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 445)
AUTHORS      Ronaldo M.F., Lennon G. and Soares M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      9704477
COMMENT      Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
```

```
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: meste@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
```

FEATURES

```
Location/Qualifiers
1..445
Source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-atk-d-06-0-UI"
/clone_lib="NIH_BMAP_M_S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)."
6:791-806, 111 g 115 t
```

BASE COUNT

```
ORIGIN
Query Match      12.9%; Score 445; DB 10; Length 445;
Best Local Similarity 100.0%; Pred. No. 3.7e-67;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 612 TGTGGAGTGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTGCTGTG 671
DB 445 TGTGGAGTGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTGCTGT 386
QY 672 TGTCAAGAGGAGGAAAAGTACTTTCATGAGTTGTTCCGGAACAGACATCAGATCTG 731
DB 385 TGTCAAGAGGAGGAAAAGTACTTTCATGAGTTGTTCCGGAACAGACATCAGATCTG 326
QY 732 TACAAATGCACTGGCGCGCAATGCACCAAGCTGTTCCACCATAGATCTAAACACCGCTCC 791
DB 325 TACAAATGCACTGGCGCGCAATGCACCAAGCTGTTCCACCATAGATCTAAACACCGCTCC 266
QY 792 TCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCCCTTGTGGATCAGGTGAA 851
DB 265 TCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCCCTTGTGGATCAGGTGAA 206
QY 852 GGCCATCCATGTGAACCATGATTCGGGCTCACCCTGGGAGCTGGAAGACAAAGCCCTGGA 911
DB 205 GGCCATCCATGTGAACCATGATTCGGGCTCACCCTGGGAGCTGGAAGACAAAGCCCTGGA 146
```

DB	175	181	235	294
DB				
Qy	CTGCTGTGATCAAGTGTGTTTTATCAATCATAGAACAATGATTCATCAGTGGGGAGG	CCATCATCGTACCAGTGGTGGCGAGGATCCCGAGAAGACCTCCACGTGTACCCCGGAGGGCC	TCATCATCATATCCCATGGTATCAGAATCCCGGAAGACCTCGGGTGTGCGTGTGAGACCC	
Db				

QY	241	CAGAGTGAAGGGACCGGTATATGACGGCCACCGTGGAGTGGCCGAGTCTGGTGGTCCATC	300
Db	295	CAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTTCATC	354

Db	355	ACACTGCAAGTGGTGGTATGCGCCACAGGGAACATTTCTGTCTCTGGGTCTTTAAGCAC	414
Qy	361	AGCTCCCTGGGCTGCCAGCCGCACTTTGATTTACAAAAACAGAGGAATCGTTTCCATGCC	420
Db	415	AGCTCCCTGAAATGCCAGCCACATTTTGATTTACAAAACAGAGGAGTAGTTTCCATGGTTC	474
Qy	421	ATCTTGAACTGACAGAGACCCAGCGAGGAGAATACCTACTCCATATTACAGCGCAACGC	480
Db	475	ATTTTGAATAATGACAGAAACCAAGCTGAGGAATACCTACTTTTATTTCAGAGTCAAGCT	534

Db	535	ACCAATTACACAATATTGTTTACAGTGAGTATAAGAAATACCTGCTTTTACACATTAGA	594
Qy	541	AGACCTTTACTTTAGSAAAGATGGAACCCAGAGATGCACCTGCTCTGCATCTCCGAGGSGTGT	600
Db	595	AGACCTTACTTTTAGAAAATGGAACCCAGAGCGCCTGGTCTGTCATATCTGAGAGCGTGT	654
Qy	601	CCGAGCGCCACTGTGGAGTGGTGCTCTGCAAGCTCCACAGGGAAAGCTGTAAAGAA	660
Db	655	CCAGCGCGATCGTGGAAATGGTGCTTTGCCATTACAGGGGGAAGCTGTAAAGAACGA	714
RESULT		9	
BQ556177			

DEFINITION	H4038E08-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone H4038E08 5', mRNA sequence.
ACCESSION	BQ556177
VERSION	BQ556177.1 GI:21457055

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1 (bases 1 to 515)
VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin
P.R., Stagg,C.A., Bassey,U., Alba,K., Hamatani,T., Kargul,G.J.,
Luo,A.G. and Ko,M.S.H.
TITLE Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
JOURNAL Unpublished (2002)
COMMENT Other_ESTs: H4038E08-3
Contact: Yong Qian

355 Cassell Drive, Suite 300, Baltimore, MD 21224-0020, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.

```

FEATURES
  High quality sequence stop: 515
  POLYA-No.
    Location/Qualifiers
      1..515
        /organism="Mus musculus"
        /strain="C57BL/6"
        /db_xref="ncbiEST:H4038E08-5"
        /db_xref="taxon:10090"

```



```
/clone="H4038E08"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/notes="vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."
BASE COUNT 139 a 138 c 128 g 110 t
ORIGIN
Query Match 12.3%; Score 426.2; DB 14; Length 515;
Best Local Similarity 92.8%; Pred. No. 6.2e-64;
Matches 478; Conservative 0; Mismatches 13; Indels 24; Gaps 2;
Oy 2725 GCTAACTCTATAAATGATTCAGAGTGGATTTAAATGGAGCGCCATTCATGCCCACA 2784
Dy 1 GCTAACTCTATAAATGATTCAGAGTGGATTTAAATGGAGCGCCATTCATGCCCACA 60
Oy 2785 GAAGGGATATCTTTGTAATCAATCCTCGCTGGGCTTTTGACTCAAGGAAGCGGCCATCC 2844
Dy 61 GAAGGGATATCTTTGTAATCAATCCTCGCTGGGCTTTTGACTCAAGGAAGCGGCCATCC 120
Oy 2845 TTCCCCCAACCTGACTTCAATTTTATAGGATGTCAGCTGGCAGGCGCAGAAAGC----- 2897
Dy 121 TTCCCCCAACCTGACTTCAATTTTATAGGATGTCAGCTGGCAGGCGCAGAAAGCGATGTAT 180
Oy 2898 -----ATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCTC 2941
Dy 181 CAGAACATGGGTGGCAAGCTCCCAACATCCATCCATCTACCAAAACAGCGGCCCTC 240
Oy 2942 AGCAGAGAGGGCGGCTCAGAGCC-CAGTCGCCACAGCGCCAGGTCGAAGATTCACAGAGAA 3000
Dy 241 AGCAGAGAGGGCAGGCTCAGAGCCGCATCGCCACAGGCGCCAGGTGAAGATTCACGGAGNA 300
Oy 3001 AGAAGTTAGCGAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCA 3060
Dy 301 AGAAGTTAGCGAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCA 360
Oy 3061 AGATTAGCTCGCCCTCTGAGGAGCGCCCTACAGCGCGTTCGCTGGAGCTTTCTCT 3120
Dy 361 AGATTAGCTCGCCCTCTGAGGAGCGCCCTACAGCGCGTTCGCTGGAGCTTTCTCT 420
Oy 3121 AGATGCTGTCTGCCATTAATCAAGTGAATCTTATAAATCAAACTCTCTCTCGCACAG 3180
Dy 421 AGATGCTGTCTGCCATTAATCAAGTGAATCTTATAAATCAAACTCTCTCTCGCACAG 480
Oy 3181 GCGGGAGAGCCCAATATGAGACTTGTGTGTGAGCC 3215
Dy 481 GTGGGAGAGCCCAATATGAGACTTGTGTGTGAGCC 515
RESULT 10
BM117950/c 484 bp mRNA linear EST 30-JAN-2002
LOCUS L0858E03-3 NIA Mouse Newborn Brain cDNA Library Mus musculus cDNA
DEFINITION clone L0858E03 3', mRNA sequence.
ACCESSION BM117950
VERSION BM117950.1 GI:17080968
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 484)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K.
and Ko, M.S.H.
Systematic Analyses of NIA Mouse Newborn Brain cDNA Library
Unpublished (2001)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
```

Email: cdnaelgsun.grc.nia.nih.gov
Plate: L0858 row: E column: 03
Seq primer: -21M13 Forward
High quality sequence stop: 484
POLYA=Yes

Location/Qualifiers
1. .484
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:L0858E03-3"
/db_xref="taxon:10090"
/clone="L0858E03"
/clone_lib="NIA Mouse Newborn Brain cDNA Library"
/tissue_type="Newborn Brain"
/dev_stage="Newborn"
/lab_host="DH10B"
/note="Organ: brain; Vector: pSPORT1 (Invitrogen); Site_1:
SalI; Site_2: NotI; Mouse cDNA project by the Laboratory
of Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://elgsun.grc.nia.nih.gov/cDNA).
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen: 5'-
PGACTAGTCTAGATCGGCGCGCGCCCTTTT-3') from 48
microgram of total RNA, treated with T4 DNA polymerase,
and purified by ethanol-precipitation. The cDNAs were
ligated to Lone-linker LL-Sal3 (Ref. Development
127:1737-1749 (2000) [PMID:10725249]), purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were digested with SalI and
NotI enzymes, and cloned into SalI/NotI site of pSPORT1
plasmid vector. The DH10B E. Coli host was transformed
with ligation mixture by the standard chemical method. The
average insert size is about 1.9 kb. The library was
constructed by Yulan Piao (NIA)."
BASE COUNT 127 a 109 c 114 g 134 t

Query Match 12.3%; Score 423; DB 13; Length 484;
Best Local Similarity 95.1%; Pred. No. 2.3e-63;
Matches 450; Conservative 0; Mismatches 15; Indels 8; Gaps 1;
Oy 2963 CCCAGTCCACAGCGCCAGGTGAAGATTCACAGAGAAAGAGTTAGCGAGGAGGCTTG 3022
Dy 474 CCCATCGCCACAGGCGCCAGGTGAAGATTCACAGAGAAAGAGTTAGCGAGGAGGCTTG 415
Oy 3023 GACCCGCGCACCTAGCAGGCTGTAGACCGCAGAGCCAGATTTAGCCTCGCCTCTCAGGA 3082
Dy 414 GACCCGCGCACCTAGCAGGCTGTAGACCGCAGAGCCAGATTTAGCCTCGCCTCTCAGGA 355
Oy 3083 AGCGCCCTACAGCGCGCTTCTCGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTACTCC 3142
Dy 354 AGCGCCCTACAGCGCGCTTCTCGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTACTCC 295
Oy 3143 AAGATGACTTCTATAAATCAAACTCTCTCCACAGCGCGGAGAGCCCAATATGAGAC 3202
Dy 294 AAGATGACTTCTATAAATCAAACTCTCTCCACAGCGCGGAGAGCCCAATATGAGAC 235
Oy 3203 TTGTTGGTGAGCGCGCTACCCCTGGGGGCGCTT-----CCACGAGCTTGGGGGAAA 3254
Dy 234 TTGTTGGTGAGCGCGCTACCCCTGGGGGCGCTTCCAGGCGCGCCCGCTTGGGGGAAA 175
Oy 3255 GCCATGATCTGAAATATAGTATATTTCTGTAATAGCTGTAACAAACACCCCGTTTT 3314
Dy 174 GCCATGATCTGAAATATAGTATATTTCTGTAATAGCTGTAACAAACACCCCGTTTT 115
Oy 3315 TTGCTAAGGAAAGCTAAATATGATTTTAAATCTATCTTTAAATCTATGTAAGTAACT 3374
Dy 114 TTGCTAAGGAAAGCTAAATATGATTTTAAATCTATCTTTAAATCTATGTAAGTAACT 55
Oy 3375 TTTTTCATCTATTAGTGATATATTTTATGATGGAATAAATCTTCTACTGTA 3427
Dy 54 TTTTTCATCTATTAGTGATATATTTTATGATGGAATAAATCTTCTACTGTA 2

```

RESULT 11
BM484050
LOCUS
DEFINITION
537756 MARC 2PTG Sus scrofa cDNA 5', mRNA linear EST 05-FEB-2002
ACCESSION
BM484050
VERSION
BM484050.1 GI:18534378
KEYWORDS
EST.
SOURCE
pig.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 507)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt.trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 2 row: D column: 23
Seq primer: ATTAGTGACACTATAG.
FEATURES
Source
Location/Qualifiers
1..507
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PTG"
/tissue_type="pooled"
/lab_host="Dhl08"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT
144 a 132 c 118 g 112 t 1 others
Query Match
Best Local Similarity 10.3%; Score 354.4; DB 13; Length 507;
Matches 425; Conservative 0; Mismatches 81; Indels 2; Gaps 2;
QY 1170 CCATGACGTGGATCTTCTCAAGCCTCATTTCTTGACACAGAGAGCCCTGGAGGA 1229
D 1 CAGGTGACGTGGACCTTCTCTC-GCAACATTTCTGTGACAAAGGGGACTCGACGA 59
QY 1230 TGGGTACAGCATATCTAAATTTTGGCATCATAGAACAAGCCAGGAGTACATATCTA 1289
D 60 CGGTACAGCATATCTAAGTTTGGCAACCATTAACACAGCAGGAGCAATACATCTTCCA 119
QY 1290 TGCA-GAAATATGATGACGCCAGTCACCAAAATGTTTCAGCGTCAATATAAGAAAGAAC 1348
D 120 TGCAGAAATATGACGACGCCAGTCACAAAGATGTTTCAGCGTCAATATAAGAAAGAAC 179
QY 1349 CTCAGTGTAGCAATATGCTTCAGCCAGCAGCGCTCTTCTCTGTATGGGTACCCGC 1408
D 180 CGCAAGTGTGCGCGAGGATCGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 239
QY 1409 TACCTCTTGGACTGGAAGAGTGTTCGACAAATCTCCCAATTCGACGAGGAATCC 1468
D 240 TACCGTCTTGGACTGGAAGAGTGTTCAGCAAGTCTCCCAATTCGACGAGGAATCA 299
QY 1469 CAGAAGGAGTTTGAATATAAAGGCTTAACAGAAAGTGTTCGACGAGTGTGTCGAGCA 1528
D 300 CGGAAGGCATCTGGAATATAAAGGCCAACAGGAAGATTTTGGACAGATGATTTCCAGCA 359

```

```

QY 1529 GTACTCTAAATATGAGTGAGCGCGGAAGGGCTTCTGGTCAAAATGCTGTGGTACAAATT 1588
D 360 GCATCTGACATAGAGGAGGCTGTCAAGGGTTTCTGGTGAAGTGTGTGCTTACAAATT 419
QY 1589 CTATGGCAGCTCTTGGGAAACCACTTTTAAACTCACCAGGCCCCCTTCCCTTTTCATCC 1648
D 420 CCCTGGGCAGCTCTTGTGAACGATCTCTTAAACTCACCAGGCCCTTCCCTTTTCATCC 479
QY 1649 AAGACACATCTCTTCTCTATGCGACCAT 1676
D 480 AGGACACATCTCTATCTATGCAACAAT 507
RESULT 12
AV713950
LOCUS
DEFINITION
AV713950 DCB Homo sapiens cDNA clone DCBADE05 5', mRNA linear EST 11-OCT-2000
ACCESSION
AV713950
VERSION
AV713950.1 GI:10795467
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 460)
Xu,X., Gu,J., Liu,F., Ou,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
,G., Cheng,Z., and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzge@hgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Source
Location/Qualifiers
1..460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DCBADE05"
/clone_lib="DCB"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/note="vector: pRipEx2; Site_1: SfiI; Site_2: SfiI"
BASE COUNT
152 a 95 c 96 g 117 t
Query Match
Best Local Similarity 9.9%; Score 340.8; DB 10; Length 460;
Matches 378; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 1906 ATGAACCCCGCCGCTATGGCATTTAGTAAACGAGGTCTCAATTCAGGTGGCGTGAAG 1965
D 1 ATGAACCCCAACAGCTTATGGAATTAGCATACAGGTGTCTCAATCCAGGTGGCGTCAAA 60
QY 1966 ATGTAAAGAGAAAGTGCAGCTGTGAAAGAAAGTCTCATGTGGAGCTCAAAATG 2025
D 61 ATGTGAAAGAAAGAGAGAGAGTGTGAAAGAGGACACTCATGTGAGAACTCAAGATG 120
QY 2026 ATGACCCACCTGGGACACCATGACAACATCTGAATCTGTGGGGGCGTGCACACTGTCA 2085
D 121 ATGACCCAGCTGGGAGGCCACGAGAATATTGTGAACCTGTGTGGGGGCGTGCACACTGTCA 180
QY 2086 GGGCCACTGTACTTGATTTTGAATATTGTTCATGTGTGATCTTCTCAACTATCTAAGA 2145
D 181 GGACCAATTTACTTGATTTTGAATACTGTGTGATCTTCTCAACTATCTAAGA 240

```

QY	2146	AGTAAAGAGAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGT	2205
Db	241	AGTAAAGAGAGAATTTTCACAGGATTTTCACAGGATTTTCACAGGATTTTCACAGGATTTTCAG	300
QY	2206	TCTTACCCCTACTTTCACAGGACATTTCAAATTCAGGATGCTGGTTCACAGGAGTTCAG	2265
Db	301	TTTACCCCACTTTCGAATCATCATCAATTTCCAGCATGCTGGTTCAGAGAGGTTTCAG	360
QY	2266	TTACACCCGCCCTTGATCATGCTCTCAGGGTTCAATGGGAATTCATTTCAATCTCTGAAGAT	2325
Db	361	ATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGAT	420
QY	2326	GAGATTGAATATGAAAACCA 2345.	
Db	421	GAAATGAATATGAAAACCA 440	
RESULT 13			
AK004947			
LOCUS			
DEFINITION		3312 bp mRNA linear HTC 19-JAN-2002	
		Mus musculus adult male liver cDNA, RIKEN full-length enriched	
		library, clone:1300008N20:colony stimulating factor 1 receptor,	
		full insert sequence.	
ACCESSION		AK004947	
VERSION		AK004947.1 GI:12836517	
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA,	
		clone_1b:RIKEN full-length enriched mouse cDNA library	
ORGANISM		Clone:1300008N20.	
		Mus musculus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE			
AUTHORS		Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE			
AUTHORS		2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
		Itoh, M., Konno, H., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to	
JOURNAL		prepare full-length cDNA libraries for rapid discovery of new genes	
MEDLINE		Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED		20499374	
REFERENCE			
AUTHORS		3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
		Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,	
		Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,	
		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,	
		Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,	
		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,	
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format	
JOURNAL		sequencing pipeline with 384 multipillar sequencer	
MEDLINE		Genome Res. 10 (11), 1757-1771 (2000)	
PUBMED		20530913	
REFERENCE			
AUTHORS		4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,	
		Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,	
		Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,	
		Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,	
		Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,	
		Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,	
		Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,	
		Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,	
		Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,	
		Baldarelli, L., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,	
		Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,	
		Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,	
		Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,	

TITLE		Functional annotation of a full-length mouse cDNA collection	
JOURNAL		Nature 409 (6821), 685-690 (2001)	
MEDLINE		21085660	
PUBMED		11217851	
REFERENCE			
AUTHORS		5 (bases 1 to 3312)	
		Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,	
		Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,	
		Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,	
		Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,	
		Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,	
		Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,	
		Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,	
		Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,	
		Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,	
		Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,	
		Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,	
		Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,	
		Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and	
		Hayashizaki, Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of	
MEDLINE		Physical and Chemical Research (RIKEN), Laboratory for Genome	
PUBMED		Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
REFERENCE		RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,	
AUTHORS		Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,	
		URL: http://genome.gsc.riken.go.jp/, Tel.81-45-503-9222,	
		Fax:81-45-503-9216)	
COMMENT		Please visit our web site (http://genome.gsc.riken.go.jp/) for	
		further details.	
		cDNA library was prepared and sequenced in Mouse Genome	
		Encyclopedia Project of Genome Exploration Research Group in RIKEN	
		Genomic Sciences Center and Genome Science Laboratory in RIKEN	
		Division of Experimental Animal Research in Riken contributed to	
		prepare mouse tissues. First strand cDNA was primed with a primer	
		[5' GAGAGAGAGCGCGCACTGAGTGTGTTTTTTTTTNN 3'], cDNA was	
		prepared by using trehalose thermo-activated reverse transcriptase	
		and subsequently enriched for full-length by cap-trapper. Second	
		strand cDNA was prepared with the primer adapter of sequence[5'	
		GAGAGAGAGAGGATCCAGAGCTCAATTAATTAACCCCCCCC 3']. cDNA was	
		cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:	
		XhoI. Host: SOLR.	
FEATURES		Location/Qualifiers	
		1..3312	
		/organism="Mus musculus"	
		/strain="C57BL/6J"	
		/db_xref="PANTOM_DB:1300008N20"	
		/db_xref="MGD:MGI:1904396"	
		/db_xref="taxon:10090"	
		/clone="1300008N20"	
		/sex="male"	
		/tissue_type="liver"	
		/clone_lib="RIKEN full-length enriched mouse cDNA library"	
		/dev_stage="adult"	
		1..3312	
		/gene="Csflr"	
		93..3026	
		/gene="Csflr"	
		/note="colony stimulating factor 1 receptor	
		data source:MGD, source key:MGI:1339758, evidence:ISS	
		putative"	
		/codon_start=1	
		/protein_id="BAB23691.1"	
		/db_xref="GI:12836518"	
		/db_xref="MGD:MGI:1339758"	
		/translation="MELGPPVLVLLATVHGGAPVIEPSPELVVEPGETVILRCVS	
		NSVEMDGPISPYWLDPPSPGSTLTTRNATPKNTCTCTELEDPMAGSTTTLXVK	
		DPAHSNLLAQEVTVVGGQEAFLPLITDTPALNKDSVLSLMREGQVLRKTVYFSPWR	

GFILRKAVLDSNTYVCKTMVNGRESTSTGIWLKVRNVHPPEPQIKLEPSKLVIRIGE
AAQIVCATNAEVFNILKRGDTKLEPLTSEQLLOEVSVGDSLLITVHADAYPSIQHY
SCVASNDVGTATMNFQVSVESVYLNTSEQLLOEVSVGDSLLITVHADAYPSIQHY
NWTYLGPFEDORKLEITQRAIYRYTEKFLFNVRKASBAGQYFLMAQNKAGNNLIF
ELTURYPEVSVTWPVNGSDVLFCDVSGVPOPSVTWMECHRGHTRDCEAQAQVWIND
THEVLVSOKPFDKVIIOQSLPIGTLKHNMTYFCRTHNSVGNSSQYFRAVSLGQSKOLF
DESLETFVYVACMSVSLVLLLLLYKYKQPKYQVWRKIIERYGNSYTFIDPTQ
LPYNKEVFPNNLOFGKTLGAGAFKVEATEFGLKEDAVLKVAVKMLKSTAHAD
KEALMSLKIMSHLGOHENIVNLGACTHGGPVLVITEYCCYGLLNFRLRKAEMILG
PSLSPGDSGDSKNIHLEKKYVRDRSDGSPSGVDYIVEMRPVSTSSSFRQDL
DKEASRLEMLDLHFSSQAQMAFLASKNCHRDVAARNVLLTSFHVAKIGDFGLA
RDIMDENVYVKVGNLFPVMAPESTFDCVTVOSDVMYSYILLWFSGLFSLGYPFG
ILVNNKFIYKVDGYQMAQVPAPEKNIYSIMOSCWLEPTRRPTFQOICFLQEQARL
ERRDQYANLPSGSGSDSGSGSSGSSSEPEESSEHLACCEPDIAQLPQNP
NYQFC"

BASE COUNT 779 a 915 c 897 g 721 t
ORIGIN

Query Match 9.7%; Score 333.6; DB 11; Length 3312;
Best Local Similarity 57.8%; Pred. No. 3e-48;
Matches 666; Conservative 0; Mismatches 459; Indels 27; Gaps 3;

QY 1697 TTGTGTTCTCATTTGTTGATCTGCCACAATACAAAACCAATTTAGGTACAGAGTC 1756
DB 1672 TGGTGCTACTGCTGTTGCTGCTTGTACAAAGTACAAGCAGAACCGAAGTACCAGTGC 1731
QY 1757 AGCTGCAGATGATCCAGGTAGTGCCTCCCTGGATAACGAGTACTTCTACCTTGACTTCA 1816
DB 1732 GCTGAAGATCATCGAG-----AGATACGAAGCAATAGCTACACCTTCATTGACCCTA 1785
QY 1817 GGGACTATGATATGACCTTAAGTGGGAGTTCCCGAGAGAGAACTTAGAGTTTGGGAAG 1876
DB 1786 CTCAGTTGCCCTACAAAGTGGAGTTCCCTCGGAACAACCTCGAGTTTGGTAGA 1845
QY 1877 TCCGTGGGCTGCGCGCTTTTCGGAGGGTGATGAACGCCACCGCTATGGCAATTAGTAAAA 1936
DB 1846 CTCCTAGGAGCGGTGCTTTGGGAAGGTGGTGGAGCTTACAGCGCTTTGGTCTGGGCAAG 1905
QY 1937 CGGAGTCTCAATTCAGGTGGCGTGAAGTCTTAAAGAGAAAGCTGACAGCTGTGAA 1996
DB 1906 AAGATGAGTGTCTGAAGGTGGCTGTGAAGTCTTAAAGTCCACGGCTCATCTGATGAGA 1965
QY 1997 AAGAAGCTCTCATCTCGAGCTCAAAATGATGACCCACCTGGGACACCATGACAAACATCG 2056
DB 1966 AGGAGGCCCTGATCTCAGAGCTGAAGATCATGATGATGATGATGATGATGATGATGATG 2025
QY 2057 TGAATCTGCTGGGGCATGACATGTCAGGGCAGTGTACTTGTATTTTGAATATTGTT 2116
DB 2026 TCAACCTCTTGGAGCGCTGACTCAGCGAGACCTGCTCTGCTCATCTGAATGCTGCT 2085
QY 2117 GCTATGTTGACCTCTCACTACCTAAGATAAGATAAGAGAGAAAGTTTTCACAGGACATGGA 2176
DB 2086 GCTATGAGACCTACTCAACTTTTCCGAAGAAAGCGCGGCTATGCTAGAGCCCAAGCC 2145
QY 2177 CAGAGATTTTAAAGGAACATAAATTTCACTTCTTACCCTTACTTTCAGGCACATCAAAAT 2236
DB 2146 TGAGTCTGTCAGA-----CTCCAGGAGGAGCTCCAGCTACAGAAACATCCACCTG 2198
QY 2237 CCAGCATGCTCTGGTTACGAGAAAGTTTCACTTACACCCCGCTTGGATCAGCTCTCAGGGT 2296
DB 2199 GAGAAGAAATATGTGCGCAGGACAGTGGCTTCTCCAGTACGAGGTGTAGACACCTACGTG 2258
QY 2297 TCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAACACAGAGAGGCTGG 2356
DB 2259 GAGATGAGCCCTGCTCGACTTCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2304
QY 2357 CAGAAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGACCTCTCTTCTGCTTTGCGT 2416
DB 2305 AAGATCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2364
QY 2417 ACCAAGTGGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2476
DB 2365 GCGAAGTGGCTCAGGAGATGCGCTTCTGCTTCTTAAACATGCAATCCACCGGACGTAG 2424

QY 2477 CAGCAGGAATCTGTTGTTGTCACCCAGGAGGTGTTGAAGATCTGTGACTTTGGACTGG 2536
DB 2425 CAGCTCGAAACGTGCTGTTGACAGCGGACATGTGCCAAGATTTGGGACTTTGGACTGG 2484
QY 2537 CCCGAGACATCTGAGGAGCTCCAGCTACGTGCTCAGGCGCAACGACGCTGCCGCTGA 2596
DB 2485 CTAGGGACATCATGAATGACTTCAACTATGTTGTCAAGGGCAATGCCCGCTGCTGTAA 2544
QY 2597 AGTGATGGCACCAGAGAGCTTATTGTAAGGATCTACACAATCAAGATGAGCTGCTGT 2656
DB 2545 AGTGATGGCACCAGAGAGCTTATTGTAAGGATCTACACAATCAAGATGAGCTGCTGT 2604
QY 2657 CCTAGGATCTCTCTCTGCGGAGATATTTCACGTGCTGAACCTTACCTTGGCATTC 2716
DB 2605 CCTAGGATCTCTCTCTGCGGAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2664
QY 2717 CTGTCAGCTAACTTCTATAAATGATTGAGTGGATTTAAAATGGAGAGCTTCT 2776
DB 2665 TAGTGAACAACAAGTTCTACAACTGTTGAAGGATGGATACCAATAGCCGCTGTAT 2724
QY 2777 ATGCCACAGAGGGATATACCTTTGTAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTG 2836
DB 2725 TTGCACCGAAGAACAATATACAGCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2784
QY 2837 GGCCATCTCTCC 2848
DB 2785 GACCACCTTCC 2796

RESULT 14

BC026713
LOCUS BC026713 5174 bp mRNA linear HTC 07-AUG-2002
DEFINITION Mus musculus, clone IMAGE:5008623, mRNA.
ACCESSION BC026713
VERSION BC026713.1 GI:19934327
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5174)
Direct Submission
Strausberg, R.
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Xoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 45 Row: b Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: no 5' EST match.

FEATURES

source

1..5174
/organism="Mus musculus"

```

/db_xref="taxon:10090"
/map="FVB/N-3"
/clone="IMAGE:5008623"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_llb="NCI_CGAP_Mam2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1356 a 1232 c 1246 g 1340 t
BASE COUNT
ORIGIN

```

Query Match 9.4%; Score 324.8; DB 11; Length 5174;
Best Local Similarity 58.0%; Pred. NO. 8.3e-47;
Matches 612; Conservative 0; Mismatches 441; Indels 3; Gaps 2;

QY	1792	AACGAGTACTCTACGTTGACTTCAGGGACTATGAATATGACCTTTAGTGGGAGTTCCCG	1851
Db	1749	AACAATTTATGTTTACATAGACCCGACGCAACTTCTCTTATGATCACAAATGGGAGTTTCCC	1808
QY	1852	AGAGAGAACTTAGAGTTTGGGAAGTCTCGGGTCTGCGCTTCGGAGGGGTGATGAAC	1911
Db	1809	AGAAACAGGCTGAGTTTGGAAAGACATTTGGGAGCTGGTGCCTTCGGGAAGTTCGTTGAG	1868
QY	1912	GCCACGCCCTATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTA	1971
Db	1869	GCCACTGCATATGGCTTGAATTAAGTCGATGCTGCGCATGACAGTTGCCGTGAAGATGCTC	1928
QY	1972	AAAGAGAAAGCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACC	2031
Db	1929	AAACCAAGTGCCCAATTTAAACAGAAAGAGAGGCCCTTAATGTGCGAACTGAAAGTCTCTGAGC	1988
QY	2032	CACCTGGGACACCATGACACATCGTCAATCTGCTGGGGCATGCACACTGTCTAGGGCCA	2091
Db	1989	TACCTGGGCAATCACATGAATATTGTGAACCTGCTTGGCGCATGCACGGTGGGAGGGCCC	2048
QY	2092	GTGTACTTGTATTTTGAATATTGTGTCTATGTTGACCTCTCAACTACCTAAAGAAGTAAA	2151
Db	2049	ACCCTGCTCATACAGAATATTGTTGCTATGTTGATCTTTTGAATTTTTCGAGAAGGAAG	2108
QY	2152	AGAGAGAAGTTTTCACAGACATGACAGAGATTTTTAAGNAACAATATTTCACTGTTCTTAC	2211
Db	2109	CGTGACTCGTTTATTTTCTCAAGCAAGAGAGACGACGAGAGGGGCACTTTATAAAGAAC	2168
QY	2212	CTTACTTTCCAGGCACATTCAAATTCACAGATGCTGTTTCACGAGAAGTTCACTTACAC	2271
Db	2169	CTTCTGCATCAACGGAGCCTTCTCTGTGACAGTTTCAATGAATATATGGACATG - AAGC	2226
QY	2272	CCGCGCTTGGATCAGCTCTCAGGGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATT	2331
Db	2227	CTGCGGTTTCTCTACGTGTCACCAAGACAGACAGAGGAGATCCGCAAGAAATAGACT	2286
QY	2332	GAATATGAANAACAGAGAGCTGGCAGAGAGAGAGAGGAGAGATTTTGAAGTCTGTCAGC	2391
Db	2287	CGTACATAGAAAGAGAGCTGACT - CCTGCCATCATGGAAGATGACAGCTGGCTCTGGAC	2345
QY	2392	TTTGAAGACCTCTCTGCTTTGCGTACCACAGTGGCCAAAGCATGGAATTCCTGGAGTTC	2451
Db	2346	CTGGATGATTGCTGAGCTTCTCTACAGGTGGCCAAAGGCATGGCGTTCCTCGCCTCC	2405
QY	2452	AAGTCGTGTGTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCACGGGAAGGTG	2511
Db	2406	AAGAATTGATTTCACAGAGATTTGGCAGCCAGGAATATCTCTCTCACTACACGGCGGATC	2465
QY	2512	GTGAAGATCTGTACTTTGGACTGGCCCGACACATCTCTGACCGACTCCAGCTACGTGCTC	2571
Db	2466	ACAAGATTTTCGATTTTCGGGCTAGCCAGAGACATCATGGAATGATTTCGAATTTACGTGCTC	2525
QY	2572	AGGGCAACGACCGCTCGCGGTGAAGTGGATGCAACCCGAGAGCTTATTTGAAGGGATC	2631
Db	2526	AAAGGAATGCAACGACTGCCGTGAAGTGGATGGCACCCAGAGAGCATTTTCAAGCTGGCGTG	2585
QY	2632	TACACAATCAAGAGTGAGCTGCTGCTCTACGGCATCTTCTCTCTGGGAGATATTTTCACTG	2691

Db	2586	TACACATTTGAAAGTAGTGTGGTCTCATGGGATTTTCTCTCGGGAGCTCTCTCCTTA	2645
Qy	2692	GGTGTGAACCCCTTACCCCTGGCATCTCTGTCGAGCGTAACCTTCATATAACATGATCAGAGT	2751
Db	2646	GGRAAGCAGCCCTTACCACGGGATCCCGTCTCAAGTTCTACACAGATGATCAAGGAA	2705
Qy	2752	GGATTTAAATGGAGCAGCCATTTCTATGCCACAGAAAGGATATACCTTTGTATGCAATCC	2811
Db	2706	GGCTTCGGATGGTTCAGCCCCGAGCAGCGCCTGCCGAAATGTATGACGTCATGAAGACT	2765
Qy	2812	TGCTGGCCTTTTGACTCAAGGAAGCGGCATCTCTTC	2847
Db	2766	TGCTGGAGCGCTGACCCCTTTGAAAGGCCAACATTC	2801

RESULT 15
 BG145808
 LOCUS
 DEFINITION
 BG145808 394 bp mRNA linear EST 01-FEB-2001
 uu98g07.yl Soares_mouse_NNGB_bcell Mus musculus cDNA clone
 IMAGE:3384737.5' similar to SW:FLT3_HUMAN P36888 FL CYTOKINE
 RECEPTOR PRECURSOR ; , mRNA sequence.
 ACCSSION
 VERSION
 BG145808
 KEYWORDS
 EST
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 394)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicapp>.
 NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ccapbs@mail.nih.gov
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1089753
 Trace considered overall poor quality
 Seq primer: -40RP from Gibco
 High quality sequence stop: 1.
 Location/Qualifiers
 1..394
 FEATURES

```

1...394
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3384757"
/clone_lib="Soares_mouse_NMGB_bcell"
/lab_host="DH10B (phage-resistant)"
(note=Organ: germinal B-cell; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCCGTGTTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaldo."
95 a 86 C 109 q 104 t

```

BASE COUNT	95 a	86 c	109 g	104 t
ORIGIN				
Query Match		9.1%	Score 314;	DB 12; Length 394;

	Query Match	9.1%	Score 314;	DB 12;	Length 394;
	Best Local Similarity	94.2%;	Pred. No. 1.7e-44;		
	Matches 326;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0;
QY	2447	AGTTCAAGTCGTGTCCACAGAGACCTGGCAGCCAGGAAATGTTGGTCACCCACGGGA	2506		
Db	48	AGTTAAAGTCGCGTGTCCACAGAGACCTGGCAGCCATGAATGTTGGTCACCCACGGGA	107		
QY	2507	AGGTGGTGAAGATCTGTGACATTTGGACTGGCCCGCAGACATCCTGAGCGACTCCAGCTACG	2566		
Db	108	AGGTGGTGAAGATCTGGACATTTGGACTGGCCCTAGACATGCTGAGCGACTACAGCTACT	167		

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:14 ; Search time 60.9693 Seconds

(without alignments)
2168.053 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLELLVLSV.....RGLRAQSPQQRVKIHRERS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:
	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	5264	100.0	992	14 AAR37502	Murine flk-2. Mus
2	5264	100.0	992	16 AAR67815	Flk2 receptor prot
3	5264	100.0	992	16 AAR67535	Mouse flk-2. Mus
4	5264	100.0	992	17 AAR97418	Human foetal liver
5	5264	100.0	992	18 AAW19874	Murine flk-2 recep
6	5264	100.0	992	20 AAY08616	Murine flk-2 prote
7	5256	99.8	992	14 AAR44994	Murine flk-2 recep
8	5256	99.8	992	14 AAR31375	Murine flk-2. Mus
9	5239	99.5	992	13 AAR28038	Murine flk-2. Mus
10	5102	96.9	1000	16 AAR81868	Flk2/flt3 tyrosine

11	4436.5	84.3	993	16 AAR75961	Human STK-1. Homo
12	4429.5	84.1	993	16 AAR67816	Flk2 receptor prot
13	4429.5	84.1	993	16 AAR67536	Human flk-2. Homo
14	4429.5	84.1	993	17 AAR97419	Murine foetal live
15	4429.5	84.1	993	18 AAW19873	Human flk-2 recep
16	4429.5	84.1	993	20 AAY08617	Human flk-2 protei
17	4426.5	84.1	993	14 AAR44995	Human flk-2. Homo
18	4421.5	84.0	993	14 AAR44995	Human flk-2 recep
19	4420.5	84.0	993	16 AAR81869	Human Flk2/flt3 ty
20	4417.5	83.9	1167	14 AAR31376	Human flk-2. Homo
21	4409.5	83.8	983	19 AAW63588	Human receptor typ
22	4408	83.7	986	19 AAW63587	Human receptor typ
23	4408	83.7	986	19 AAW63589	Human receptor typ
24	4404	83.7	994	19 AAW63586	Human receptor typ
25	2806.5	53.3	665	19 AAW63585	Human receptor typ
26	2373.5	45.1	749	19 AAW78002	Protein PHON32390.
27	2226.5	42.3	481	15 AAR47579	Soluble Flk-2. Mu
28	1223.5	23.2	972	23 AAU11935	Colony stimulating
29	1223	23.2	975	22 AAE07148	Murine Kit/stem ce
30	1223	23.2	975	22 AAE07148	Mutant murine Kit/
31	1221.5	23.2	972	23 AAU79039	Human macrophage c
32	1221.5	23.2	972	23 AAU11936	Colony stimulating
33	1221.5	23.2	972	23 AAU11941	Colony stimulating
34	1221.5	23.2	972	23 AAU73585	Colony stimulating
35	1221.5	23.2	1055	22 ABG15479	Novel human diagno
36	1220.5	23.2	972	23 AAU11934	Colony stimulating
37	1220.5	23.2	972	23 AAU11937	Colony stimulating
38	1219.5	23.2	972	23 AAU11940	Colony stimulating
39	1216.5	23.1	972	23 AAU11938	Colony stimulating
40	1215	23.1	977	21 AAV51322	Bovine c-kit bk-1
41	1212.5	23.0	972	23 AAU11939	Colony stimulating
42	1197.5	22.7	2129	22 ABG15478	Novel human diagno
43	1197.5	22.7	2129	22 ABG20749	Novel human diagno
44	1196	22.7	976	22 AAE07145	Human Kit/stem cel
45	1196	22.7	976	22 AAU00375	Human stem cell gr

ALIGNMENTS

RESULT 1

AAR37502
ID AAR37502 standard; Protein; 992 AA.

AC AAR37502;

DT 19-OCT-1993 (first entry)

DE Murine flk-2.

XX Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;
KW hematopoietic cell; mature; family; conserved; region;
KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;
KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;
KW hierarchy; transduction; T-lymphoid; lineage.

XX Mus musculus.

XX Key Location/Qualifiers

FT Peptide 1..27
FT Protein /note= "Hydrophobic leader sequence"

FT Domain /note= "Mature murine flk-2"

FT Region /note= "Extracellular receptor domain"

FT Domain /note= "Transmembrane region"

FT Domain /note= "Intracellular catalytic domain"

FT Domain /note= "Catalytic sub-domain"

FT Domain /note= "Catalytic sub-domain"

Domain 832..834
/note= "Catalytic sub-domain"
FT FT
Domain 857..862
/note= "Catalytic sub-domain"
FT FT
Domain 872..878
/note= "Catalytic sub-domain"
FT FT
Region 736..812
/note= "flk-2 signature sequence"
FT FT
XX
W09310136-A.
PN
XX
27-MAY-1993.
XX
16-NOV-1992; 92WO-US09893.
XX
15-NOV-1991; 91US-0793065.
XX
(UYPR-) UNIV PRINCETON.
PA
XX
Lemischka IR;
PI
XX
WPI; 1993-182479/22.
DR
N-PSDB; AAQ40914.
DR

XX
Totipotent haematopoietic stem cell receptors, their ligands and
PT DNA sequences - for treating anaemia(s) and bone marrow damage
PT due to e.g. cancer chemotherapy or radiotherapy
XX
XX
PS Claim 37; Fig 1a; 127pp; English.
XX

CC This sequence represents the murine receptor protein tyrosine kinase
CC (pTK), flk-2. The nucleic acid encoding this receptor is expressed
CC in primitive hematopoietic cells and not in mature hematopoietic
CC cells. Members of this family of pTK's can be recognised by the
CC conserved amino acid regions in the catalytic domain. This family
CC of pTK's also contains c-kit. These new receptors are termed fetal
CC liver kinases (flk's) after the tissue in which they were discovered.
CC flk-2 is also expressed in fetal spleen, fetal thymus, adult brain
CC and adult bone marrow. flk-2 is expressed in individual multipotential
CC CFU-Blast colonies capable of generating numerous multilineage colonies
CC upon replating. It is likely therefore, that flk-2 is expressed in
CC the entire primitive portion of the hematopoietic hierarchy. This is
CC consistent with flk-2 being important in transducing putative self-
CC renewal signals from the environment. flk-2 is the first receptor
CC pTK known to be expressed in the T-lymphoid lineage.
XX
XX Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 14; Length 992;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRLLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60
DB 1 MRALQSRDRLLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60
QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWFKHSSLCQOPHD 120
DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWFKHSSLCQOPHD 120
QY 121 LQNRGIVSMALLNTVETQAGEYLLHIOSEYANYTVLFTVNVYRDQYVLRPRYFRKMENQ 180
DB 121 LQNRGIVSMALLNTVETQAGEYLLHIOSEYANYTVLFTVNVYRDQYVLRPRYFRKMENQ 180
QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEAGPVAVVRKEEVLHFLGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEAGPVAVVRKEEVLHFLGTDIRCCARNALGR 240
QY 241 ECTLFTIDLNQAPQSTLPQLFLKVGPEPLWIRCAIHVNHGFGLTWELEKALBEGSYFE 300
DB 241 ECTLFTIDLNQAPQSTLPQLFLKVGPEPLWIRCAIHVNHGFGLTWELEKALBEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPQSALVTILEKGFINATSSQEEY 360

DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCFSVRKAYPRICTWIFSOASPCPEORGLDGYISIKFCDHKNKPGEYIFY 420
DB 361 EIDPYEKFCFSVRKAYPRICTWIFSOASPCPEORGLDGYISIKFCDHKNKPGEYIFY 420
QY 421 AENDDAQFTKMTNIRKKPOVLANASQASQSSDGYPLPSWTWKKCDKSPNCTEELP 480
DB 421 AENDDAQFTKMTNIRKKPOVLANASQASQSSDGYPLPSWTWKKCDKSPNCTEELP 480
QY 481 EGVNKKANRVFGQWVSSTLNMSSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
DB 481 EGVNKKANRVFGQWVSSTLNMSSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
QY 541 DNISFYATIGLCLPFIVVLIVLICHYKKQFYESOLOMIOVTGPLDNEYFYVDRDYEY 600
DB 541 DNISFYATIGLCLPFIVVLIVLICHYKKQFYESOLOMIOVTGPLDNEYFYVDRDYEY 600
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720
DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720
QY 721 EBNFSSYPTFOAHNSNMPGSRREVOLHPDLQSLGFNGNSIHSEDEIEYENOKRLAEDEE 780
DB 721 EBNFSSYPTFOAHNSNMPGSRREVOLHPDLQSLGFNGNSIHSEDEIEYENOKRLAEDEE 780
QY 781 EDLNVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVICDFGLARDIL 840
DB 781 EDLNVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVICDFGLARDIL 840
QY 841 SDSYVVRGNARLPVKWMAPELSLEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
DB 841 SDSYVVRGNARLPVKWMAPELSLEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEOPFAYTEGIVFMQSCWAFDRSRKPSFPNLTSLGCOLAEAEACIRT 960
DB 901 FYKLIOGFKMEOPFAYTEGIVFMQSCWAFDRSRKPSFPNLTSLGCOLAEAEACIRT 960
QY 961 SIHLPKQAAPQGRGLRAQSPQRQVKIHRERS 992
DB 961 SIHLPKQAAPQGRGLRAQSPQRQVKIHRERS 992
RESULT 2
AAR67815
ID AAR67815 standard; Protein; 992 AA.
XX
AC AAR67815;
XX
DT 16-AUG-1995 (first entry)
XX
DE Flk2 receptor protein-tyrosine-kinase.
XX
KW Mouse Flk2; receptor protein-tyrosine-kinase; primitive
KW hematopoietic cell; fetal liver kinase; diagnostic ligand
KW isolation; bone marrow disease therapy.
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /note= "signal peptide"
FT Protein 28..992
FT /note= "mature protein"
FT Domain 28..544
FT /note= "extracellular receptor domain"
FT Domain 545..564
FT /note= "transmembrane region"

FT Domain 565..992 /note= "intracellular catalytic domain" 360
FT Domain 618..623 /note= "catalytic sub-domain" 420
FT Domain 811..819 /note= "catalytic sub-domain" 420
FT Domain 832..834 /note= "catalytic sub-domain" 480
FT Domain 857..862 /note= "catalytic sub-domain" 480
FT Domain 872..878 /note= "catalytic sub-domain" 540
FT Region 709..785 /note= "signature sequence typical of Flk2" 540
XX W09500554-A.
XX 05-JAN-1995.
XX 17-JUN-1994; 94WO-US06944.
XX 18-JUN-1993; 93US-0080244.
XX 21-JUN-1993; 93US-0081508.
XX 23-NOV-1993; 93US-0157490.
XX (UYP-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1995-052014/07.
XX DR N-PSDB; AAQ81012.
XX
XX Ligand for receptor protein tyrosine kinase - useful for the
XX stimulation of primitive haematopoietic stem cells causing
XX proliferation and/or differentiation
XX
XX Disclosure; Fig 1a; 131pp; English.
XX
XX The sequence corresponds to a mouse Flk2 (fetal liver kinase)
XX receptor protein-tyrosine-kinase, which is expressed in primitive
XX hematopoietic cells but not in mature hematopoietic cells. The
XX protein is useful in isolation of receptor ligands, which have
XX applications in diagnosis of bone marrow disorders and in
XX stimulating proliferation and/or differentiation of primitive
XX hematopoietic stem cells.
XX
SQ Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 16; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALQSDRRLLLVLSVMIETVTDLPVIRKVLISHENNGSSACKPSSRYMVRGS 60
Db 1 MRALQSDRRLLLVLSVMIETVTDLPVIRKVLISHENNGSSACKPSSRYMVRGS 60
Qy 61 PEDLOCTPRRQSEGVYEAATVEVAESGSIITLQVQLATPGDLSCLWVFKHSLGQCPHFD 120
Db 61 PEDLOCTPRRQSEGVYEAATVEVAESGSIITLQVQLATPGDLSCLWVFKHSLGQCPHFD 120
Qy 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRRPFRKMNQ 180
Db 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRRPFRKMNQ 180
Qy 181 DALLCISGVPEPTVEWVLCSSHRESCKEFGPAVVRKEKVLHFGTDIRCCARNALGR 240
Db 181 DALLCISGVPEPTVEWVLCSSHRESCKEFGPAVVRKEKVLHFGTDIRCCARNALGR 240
Qy 241 ECTKLFTIDLNOAQOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEESYFE 300
Db 241 ECTKLFTIDLNOAQOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEESYFE 300
Qy 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCSKKHPSQSALVTILEKGINATSSQEEY 360

Db 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCSKKHPSQSALVTILEKGINATSSQEEY 360
Qy 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQEQRLGEGYSISKFCDHKNKPGEIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQEQRLGEGYSISKFCDHKNKPGEIFY 420
Qy 421 AENDDAQFTKMTLNIRKKPOVLANASASQSSDGYPLPSWTWKKCSKSPNCTEIP 480
Db 421 AENDDAQFTKMTLNIRKKPOVLANASASQSSDGYPLPSWTWKKCSKSPNCTEIP 480
Qy 481 EGYWKKKANRKYFGQWVSSTLNMSKAGLLVKCCAYNSMGTSCTEIFLNSPGPFPIQ 540
Db 481 EGYWKKKANRKYFGQWVSSTLNMSKAGLLVKCCAYNSMGTSCTEIFLNSPGPFPIQ 540
Qy 541 DNISFYATIGLCLPFTIVLVILICHYKKOFRYESQLOMIOVTGPDNNEYFYVDFRDEY 600
Db 541 DNISFYATIGLCLPFTIVLVILICHYKKOFRYESQLOMIOVTGPDNNEYFYVDFRDEY 600
Qy 601 DLKWEFPRENLEFGKVLGSCAFGRVMNATAYGISKTCVSIQAVKMLKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSCAFGRVMNATAYGISKTCVSIQAVKMLKADSCKEALM 660
Qy 661 SELKMMTHLGHHDNIYNLLGACTLSGVPVLIFFEYCCYGDLLNLYRSKRKFHRTWTIEFK 720
Db 661 SELKMMTHLGHHDNIYNLLGACTLSGVPVLIFFEYCCYGDLLNLYRSKRKFHRTWTIEFK 720
Qy 721 EHNFSYPTFOAHSNSMPGSRREVQLHPDLQSLGFGNGNSIHSEDEIEYENQKRLAEEEE 780
Db 721 EHNFSYPTFOAHSNSMPGSRREVQLHPDLQSLGFGNGNSIHSEDEIEYENQKRLAEEEE 780
Qy 781 EDLNVLTFFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLTGKVKVYKICDFGLARDIL 840
Db 781 EDLNVLTFFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLTGKVKVYKICDFGLARDIL 840
Qy 841 SDSSYVVRGNARLPVKWMAPESLFEGITTKSDVWSYGILLWEIFSLGVNPPGIPVDAN 900
Db 841 SDSSYVVRGNARLPVKWMAPESLFEGITTKSDVWSYGILLWEIFSLGVNPPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPFYATEGIYFVMQSCWAFDSKRKPSFNLTSFLGCQLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATEGIYFVMQSCWAFDSKRKPSFNLTSFLGCQLAEAEACIRT 960
Qy 961 SIHLPKQAAPQQRGGLRAQSPQVQKIHRSRS 992
Db 961 SIHLPKQAAPQQRGGLRAQSPQVQKIHRSRS 992
RESULT 3
AAR67535
ID AAR67535 standard; Protein; 992 AA.
XX
AC AAR67535;
XX
DT 04-JUL-1995 (first entry)
XX
DE Mouse flk-2.
XX
KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Domain 28..544 /label= Sig_peptide
FT Region 545..564 /label= Extracellular_receptor_domain
FT Domain 565..992 /label= Transmembrane_region
FT Domain 992 /label= Intracellular_catalytic_domain
XX

PR 28-JUN-1991; 910S-0728913.
 PR 15-NOV-1991; 910S-0793065.
 PR 24-DEC-1991; 910S-0813593.
 PR 26-JUN-1992; 920S-0906397.
 PR 12-NOV-1992; 920S-0975049.
 PR 30-APR-1993; 930S-0055269.
 PR 31-OCT-1994; 940S-0252517.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 XX Lemischka IR;
 XX
 DR WPI; 1996-392678/39.
 DR N-PSDB; AAT38733.
 XX
 XX
 PT Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,
 PT for isolating haematopoietic stem cells expressing receptor and for
 PT obtaining ligands
 XX
 PS Claim 1; Columns 27-34; 50pp; English.
 XX
 CC The present sequence is human foetal liver kinase 2 (flk-2), a
 CC protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-2 can be used to
 CC assay for flk receptors on the surface of primitive haematopoietic
 CC stem cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 XX
 SQ Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 17; Length 992;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRDRRLLLVLSVMIETVNDLPVICKVLISHENNGSSAGKPSRYVRGS 60
 DB 1 MRALAQRDRRLLLVLSVMIETVNDLPVICKVLISHENNGSSAGKPSRYVRGS 60
 QY 61 PEDLQTPRRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLGCPHFD 120
 DB 61 PEDLQTPRRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLGCPHFD 120
 QY 121 LQNRGIVSMALNVTETQAGEYLLHQSERANYTVLFTVNVDRDTQYVLRPRYFRKMENQ 180
 DB 121 LQNRGIVSMALNVTETQAGEYLLHQSERANYTVLFTVNVDRDTQYVLRPRYFRKMENQ 180
 QY 181 DALLCISEGVPEPTVEMVLCSSHRECKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
 DB 181 DALLCISEGVPEPTVEMVLCSSHRECKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
 QY 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKATHVNHGFLTWELDKALEEGSYFE 300
 DB 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKATHVNHGFLTWELDKALEEGSYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSVGRNDTGYTSCSSKHPSQSALVTILEKGFINATSSQEEY 360
 DB 301 MSTYSTNRTMIRILLAFVSVGRNDTGYTSCSSKHPSQSALVTILEKGFINATSSQEEY 360
 QY 361 EIDPYEKFCFSVRFKAYPIRITWIFSQAFCEQRLGDDGYISKFCDHKNKPGYIFY 420
 DB 361 EIDPYEKFCFSVRFKAYPIRITWIFSQAFCEQRLGDDGYISKFCDHKNKPGYIFY 420
 QY 421 AENDDAQFTKMTFLNLRKPKQVLANASASQASCSGGYPLPSWTWKCKSDKSPNCTEEP 480
 DB 421 AENDDAQFTKMTFLNLRKPKQVLANASASQASCSGGYPLPSWTWKCKSDKSPNCTEEP 480
 QY 481 EGVNKNKANKRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540
 DB 481 EGVNKNKANKRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540

QY 541 DNISFYATIGLCLPFIVVLIVLIVLICHYKKQFYESQLQMIOVTGPDONEYFYVDFROYEY 600
 DB 541 DNISFYATIGLCLPFIVVLIVLIVLICHYKKQFYESQLQMIOVTGPDONEYFYVDFROYEY 600
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
 QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPPVYLIFECYCCGDDLNLNLSRKREKPHRTWTEIFK 720
 DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPPVYLIFECYCCGDDLNLNLSRKREKPHRTWTEIFK 720
 QY 721 EHNFSYPTFOAHNSNSMPGSRVQLHPPLDQLSGFNNGNSIHSEDEIEYENOKRLAEDEE 780
 DB 721 EHNFSYPTFOAHNSNSMPGSRVQLHPPLDQLSGFNNGNSIHSEDEIEYENOKRLAEDEE 780
 QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 QY 841 SDSYVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
 DB 841 SDSYVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
 QY 901 FYKLIOGGRKMEOPFYATEGIIYFVQMCSWAFDSRKRSPFNLTSLFGLCOLAEAEACIRT 960
 DB 901 FYKLIOGGRKMEOPFYATEGIIYFVQMCSWAFDSRKRSPFNLTSLFGLCOLAEAEACIRT 960
 QY 961 SIHLPKQAAPQQRGGLRAQSPQROVKIHRERS 992
 DB 961 SIHLPKQAAPQQRGGLRAQSPQROVKIHRERS 992

RESULT 5
 AAW19874
 ID AAW19874 standard; Protein; 992 AA.
 XX
 AC AAW19874;
 XX
 DT 19-AUG-1997 (first entry)
 XX
 DE Murine flk-2 receptor.
 XX
 DE Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
 KW PKT; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.
 XX
 OS Mus musculus.
 XX
 FH Key
 FT Peptide
 FT 1..27
 FT /note= "Signal peptide"
 FT Domain
 FT 28..544
 FT /note= "Extracellular receptor domain"
 FT Region
 FT 545..564
 FT /note= "transmembrane region"
 FT Domain
 FT 565..992
 FT /note= "intracellular catalytic domain"
 FT Active-site
 FT 618..623
 FT /note= "Catalytic sub-domain"
 FT Active-site
 FT 811..819
 FT /note= "Catalytic sub-domain"
 FT Active-site
 FT 832..834
 FT /note= "Catalytic sub-domain"
 FT Active-site
 FT 857..862
 FT /note= "Catalytic sub-domain"
 FT Active-site
 FT 872..878
 FT /note= "Catalytic sub-domain"
 FT Peptide
 FT 709..785
 FT /note= "Signature sequence characteristic of flk-2"
 XX

PN US5621090-A.
XX 15-APR-1997.
XX 02-APR-1991; 91US-0679666.
XX 26-JUN-1992; 92US-0906397.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
XX (UYPR-) UNIV PRINCETON.
PA Lemischka IR;
PI WPI; 1997-235228/21.
XX N-PSDB; AAT72118.
DR
XX
PT Protein containing the extracellular domain of human flk-2 - used
PT for identification of primitive haematopoietic cell proliferation
PT and differentiation stimulatory ligands, e.g. for treating anaemia
XX
XX Disclosure; Fig 1A; 55pp; English.
PS
XX This sequence represents murine fetal liver kinase 2 (flk2). flk-2 is
CC a receptor protein tyrosine kinase (ptk) and is important in transducing
CC putative self-renewal signals from the environment. flk-2 is expressed
CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
CC and it is thought that flk-2 is expressed in the entire primitive portion
CC of the haematopoietic hierarchy. The invention concerns a recombinant
CC nucleic acid, preferably mRNA, which encodes a protein containing only
CC the extracellular domain of human flk-2 and lacking the flk-2 intra-
CC cellular catalytic domain. The resultant protein represents a soluble
CC form of flk-2 which is used to isolate specific ligands for flk-2. These
CC ligands can be used to stimulate proliferation and/or differentiation of
CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused
CC by cancer treatment or radiation.
XX
XX Sequence 992 AA;
SQ

Query Match 100.0%; Score 5264; DB 18; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVTKCVLISHENNGSSAGKPSRYMRVRS 60
Db 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVTKCVLISHENNGSSAGKPSRYMRVRS 60
QY 61 PEDLOCTPRQSEGTVEATVEAESGSLTQVQLATPGDLSCLVWFKHSSLCQPHFD 120
Db 61 PEDLOCTPRQSEGTVEATVEAESGSLTQVQLATPGDLSCLVWFKHSSLCQPHFD 120
QY 121 LQNRGIVSMALLNVETQAGEYLLHIOSEANRYTVLTNNVDTQVLRPRPYKRMNQ 180
Db 121 LQNRGIVSMALLNVETQAGEYLLHIOSEANRYTVLTNNVDTQVLRPRPYKRMNQ 180
QY 181 DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEVHLFETDIRCCARNALGR 240
Db 181 DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEVHLFETDIRCCARNALGR 240
QY 241 ECTKLFTIDNQAPQSTLPQLFLKVGPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
Db 241 ECTKLFTIDNQAPQSTLPQLFLKVGPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEFCFSVRKAYPRIRCTWIFSOAPFCBQRLGDCYSISKFDHKNKPGEVIFY 420
Db 361 EIDPYEFCFSVRKAYPRIRCTWIFSOAPFCBQRLGDCYSISKFDHKNKPGEVIFY 420

QY 421 AENDDAQFTKMFNTLNIRKKQVLANASASQSSSDGYPLPSWTWKKCSOKSPNCTBEIP 480
Db 421 AENDDAQFTKMFNTLNIRKKQVLANASASQSSSDGYPLPSWTWKKCSOKSPNCTBEIP 480
QY 481 EGYMKNKANKRVFGQWVSSSTLNSEAGKLLVKCCAYNSMGTSCTETIFLNSPGFFPIQ 540
Db 481 EGYMKNKANKRVFGQWVSSSTLNSEAGKLLVKCCAYNSMGTSCTETIFLNSPGFFPIQ 540
QY 541 DNISFYATIGLCLPFIVVLIVLICHKYKQFRYESQLOMQVOTGPLDNEYFYVDFRDYEY 600
Db 541 DNISFYATIGLCLPFIVVLIVLICHKYKQFRYESQLOMQVOTGPLDNEYFYVDFRDYEY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGYSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGYSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHNDINVLNLLGACTLSGVPYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720
Db 661 SELKMMTHLGHNDINVLNLLGACTLSGVPYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720
QY 721 EHNFSYPTQAHNSNSMPCGSREVQLHPPDLQSLGFCNGNSIHSEDEIYEYENOKRLAEDEE 780
Db 721 EHNFSYPTQAHNSNSMPCGSREVQLHPPDLQSLGFCNGNSIHSEDEIYEYENOKRLAEDEE 780
QY 781 EDNLNLTFFEDLLCFAYOVAKMEFLEKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 781 EDNLNLTFFEDLLCFAYOVAKMEFLEKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
QY 841 SDSYVVYVGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVVDAN 900
Db 841 SDSYVVYVGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVVDAN 900
QY 901 FYKLIQSGFMQEPFYATEGIYFVMSQWAPDSRKRSFPNLTSLFGCLAEAEACIRT 960
Db 901 FYKLIQSGFMQEPFYATEGIYFVMSQWAPDSRKRSFPNLTSLFGCLAEAEACIRT 960
QY 961 SIHLPKOAAPOORGLRAQSPOROVKIHRRS 992
Db 961 SIHLPKOAAPOORGLRAQSPOROVKIHRRS 992

RESULT 6
AAY08616
ID AAY08616 standard; Protein; 992 AA.
XX AAY08616;
XX 05-AUG-1999 (first entry)
XX Murine flk-2 protein.
XX Murine, flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
KW monoclonal; polyclonal; antibody; tyrosine kinase.
XX Mus sp.
XX US5912133-A.
XX 15-JUN-1999.
XX 10-FEB-1998; 98US-0021324.
XX 19-NOV-1992; 92US-0977451.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252498.
PR 15-FEB-1996; 96US-0601891.

XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1999-357194/30.
XX N-PSDB; AAX77514.
XX Isolating hematopoietic cells expressing fetal liver kinase 1
XX receptors
XX Disclosure; Fig 1a; 59pp; English.
XX This invention describes a novel method of isolating cells expressing
XX fetal liver kinase 1 (flk-1) receptors on their surface and comprises
XX binding the cells to a polyclonal or monoclonal antibody specific to
XX the Flk-1 receptor and isolating the cells that have bound to the
XX antibody. The method can be used to isolate hematopoietic stem cells in
XX any mammal but preferably a rat, mouse, rabbit or human. The proteins of
XX the invention belong to the receptor protein family. This sequence
XX represents the murine flk-2 protein which is used in the method of
XX the invention.
XX
XX
SQ Sequence 992 AA;
Query Match 100.0%; Score 5264; DB 20; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALAQRSDRRLLLVLSVMILETVTQDLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60
Db 1 MRALAQRSDRRLLLVLSVMILETVTQDLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60
Qy 61 PEDLOCTPRROSEGVYEEAATVEVAESGSLTQVOLATPGDLSCLVFKHSLGCGPHFD 120
Db 61 PEDLOCTPRROSEGVYEEAATVEVAESGSLTQVOLATPGDLSCLVFKHSLGCGPHFD 120
Qy 121 LONRGIVSMALNVTETQAGEYLLHQSERANYTLFTVNVREDTOLYVLRFRYFRKMNQ 180
Db 121 LONRGIVSMALNVTETQAGEYLLHQSERANYTLFTVNVREDTOLYVLRFRYFRKMNQ 180
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Qy 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
Db 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
Qy 301 MSTYSTNTMTIRILLAFVSSVGRNDTGYYTCSKHPQSALVTILEKGFINATSSOEY 360
Db 301 MSTYSTNTMTIRILLAFVSSVGRNDTGYYTCSKHPQSALVTILEKGFINATSSOEY 360
Qy 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOAFSPCEORGLDGYISKFCDHKNKPGEYIF 420
Db 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOAFSPCEORGLDGYISKFCDHKNKPGEYIF 420
Qy 421 AENDDAQFTKMTLNIKKPQVLANASQASCSDDGYPLPSWTKKCDKSPNCTEIP 480
Db 421 AENDDAQFTKMTLNIKKPQVLANASQASCSDDGYPLPSWTKKCDKSPNCTEIP 480
Qy 481 EGVWKNKANRKYFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGVWKNKANRKYFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Qy 541 DNISFYATIGLCPLFTVILVILICHYKQFRYESOLOMIQVTPGLDNEYFYVDRDVEY 600
Db 541 DNISFYATIGLCPLFTVILVILICHYKQFRYESOLOMIQVTPGLDNEYFYVDRDVEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNMNATAYIGISKTGVSIOQAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNMNATAYIGISKTGVSIOQAVKMLKEKADSCKEALM 660

Qy 661 SELKMTHTLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLRSKREKPHRTWTETFK 720
Db 661 SELKMTHTLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLRSKREKPHRTWTETFK 720
Qy 721 EHNFSYPTFOAHNSMPPGSRREVOLHPPLDQLSGFGNGSIHSEDEIEYENQKRLAEE 780
Db 721 EHNFSYPTFOAHNSMPPGSRREVOLHPPLDQLSGFGNGSIHSEDEIEYENQKRLAEE 780
Qy 781 EDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPELSIFEGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPELSIFEGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLIOGFKMQPYATEGIYFVMSQWAFDSRKPSFPNLTSLGCGQLAEAEACIRT 960
Db 901 FYKLIOGFKMQPYATEGIYFVMSQWAFDSRKPSFPNLTSLGCGQLAEAEACIRT 960
Qy 961 SIHLPKQAAPQORGLRAOSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQORGLRAOSPORQVKIHRERS 992
RESULT 7
AAR44994
ID AAR44994 standard; Protein; 992 AA.
XX
XX AAR44994;
XX AC
XX 27-JUN-1994 (first entry)
XX
DE Murine flk-2 receptor protein tyrosine kinase.
XX
XX Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
KW primitive; totipotent; haematopoietic cell; stem cell; proliferation;
KW mfk; stromal cell.
XX
OS Mus musculus.
XX
Key Location/Qualifiers
FH 1..27
FT Peptide
FT /label= signal_peptide
FT /note= "hydrophobic leader"
FT Protein
FT 28..992
FT /label= flk-2
FT /note= "mature protein"
FT Domain
FT 28..544
FT /label= extracellular_domain
FT Region
FT 545..564
FT /label= transmembrane_region
FT Domain
FT 565..992
FT /label= intracellular_catalytic_domain
FT Domain
FT 618..623
FT /note= "catalytic subdomain"
FT Domain
FT 811..819
FT /note= "catalytic subdomain"
FT Domain
FT 832..834
FT /note= "catalytic subdomain"
FT Domain
FT 857..862
FT /note= "catalytic subdomain"
FT Domain
FT 872..878
FT /note= "catalytic subdomain"
XX
XX US5270458-A.
PN
XX
XX 14-DEC-1993.
PD
XX
XX 02-APR-1991; 91US-0679666.
PF
XX
XX 02-APR-1991; 91US-0679666.
PR
XX 28-JUN-1991; 91US-0728913.
PR

XX PF 26-JUN-1992; 92WO-US05401.
XX PR 28-JUN-1991; 91US-0728913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PR 02-APR-1992; 92WO-US02750.
XX PA (UUPR-) UNIV PRINCETON.
XX PI Lemischka IR;
XX DR WPI; 1993-036323/04.
XX DR N-PSDB; AAQ45247.
XX XX Nucleic acid encoding receptor protein tyrosine kinase - allows
PT development of ligands to stimulate proliferation and/or
PT differentiation of mammalian haematopoietic stem cells
XX PS Claim 5; Fig 1a; 78pp; English.
XX CC This sequence represents a murine receptor protein tyrosine kinase
CC which belongs to a new functional class of protein tyrosine kinases
CC (ptks). ptks in this class are expressed in primitive mammalian
CC hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC).
CC This protein is an example of a receptor ptk and is called fetal liver
CC kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and
CC thymus, and adult brain and marrow. Expression of flk-2 mRNA occurs
CC in the most primitive thymocyte subset, which is believed to be
CC uncommitted. Therefore, thymocytes expressing flk-2 may be multi-
CC potential. flk-2 is the first receptor tyrosine kinase known to be
CC expressed in the T-lymphoid lineage.
XX XX
SQ Sequence 992 AA;

Query Match 99.8%; Score 5256; DB 14; Length 992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRRLLLVLSVILETVNODLPVICKVLISHENNGSSAGKSSYRMVRGS 60
Db 1 MRALAQRSDRRLLLVLSVILETVNODLPVICKVLISHENNGSSAGKSSYRMVRGS 60
Qy 61 PEDLOCTPRROSEGVYEAATVEVAESGSLTQVLATPGDLSCLWVFKHSSIGCOPHPD 120
Db 61 PEDLOCTPRROSEGVYEAATVEVAESGSLTQVLATPGDLSCLWVFKHSSIGCOPHPD 120
Qy 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLTVNVRTQLYVLRPPYFRKMNQ 180
Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLTVNVRTQLYVLRPPYFRKMNQ 180
Qy 181 DALLCISGVPEPTVWVLCSSHRSCKEEGPAVVRKEKVLHFLGTDIRCCARNALGR 240
Db 181 DALLCISGVPEPTVWVLCSSHRSCKEEGPAVVRKEKVLHFLGTDIRCCARNALGR 240
Qy 241 ECTKLFTIDLQAPOSTLPOLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYPE 300
Db 241 ECTKLFTIDLQAPOSTLPOLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYPE 300
Qy 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360
Qy 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQRLGLEDGYISIKFCDHKNKPGEIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQRLGLEDGYISIKFCDHKNKPGEIFY 420
Qy 421 AENDDAQFTKMTLNIRKPKQVLANASQSCSDGYPPLPSWTWKKCSDKSPNCTEIP 480
Db 421 AENDDAQFTKMTLNIRKPKQVLANASQSCSDGYPPLPSWTWKKCSDKSANCTEIP 480
Qy 481 EGVWNKANKRVFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPFPFIQ 540
Db 481 EGVWNKANKRVFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPFPFIQ 540

Db 481 EGVWNKANKRVFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPFPFIQ 540
Qy 541 DNISYATIGLCLPFIWLVILVILCHIKYKKQFRYESQLMQIQTGPDLDNEYFYVDFRDEY 600
Db 541 DNISYATIGLCLPFIWLVILVILCHIKYKKQFRYESQLMQIQTGPDLDNEYFYVDFRDEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQAVKMLKEKADSCKEALM 660
Qy 661 SELKMMTHLGHHDNIIVNLGACTLSGVPVLIIFYCYGDLNLYLSRKREKFRHTWTBIFK 720
Db 661 SELKMMTHLGHHDNIIVNLGACTLSGVPVLIIFYCYGDLNLYLSRKREKFRHTWTBIFK 720
Qy 721 EHNFSYPTFOAHNSMPSGSRREVQLHPDQLDLSGNGNSIHSEDETEIYENOKRLAEEEE 780
Db 721 EHNFSYPTFOAHNSMPSGSRREVQLHPDQLDLSGNGNSIHSEDETEIYENOKRLAEEEE 780
Qy 781 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVKVICDFGLARDIL 840
Db 781 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVKVICDFGLARDIL 840
Qy 841 SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
Db 841 SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
Qy 901 FYKLIQSGFKMQPFPYATGEGIVFMQSWAFDSRKRSPFNLTSLGCGLAEEACIRT 960
Db 901 FYKLIQSGFKMQPFPYATGEGIVFMQSWAFDSRKRSPFNLTSLGCGLAEEACIRT 960
Qy 961 SIHLPKQAAPOQRGGLRAOSPORQVXIHRERS 992
Db 961 SIHLPKQAAPOQRGGLRAOSPORQVXIHRERS 992

RESULT 9
AAR28038
ID AAR28038 standard; Protein; 992 AA.
XX AC AAR28038;
XX DT 15-MAR-1993 (first entry)
XX DE Murine flk-2.
XX KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;
XX OS Mus musculus.
XX FH Key Location/Qualifiers
FT Peptide 1..27 /note= "hydrophobic leader"
FT Domain 28..544 /note= "extracellular receptor domain"
FT Region 545..564 /note= "transmembrane region"
FT Domain 565..992 /note= "intracellular catalytic domain"
FT Region 709..785 /note= "signature sequence of flk-2"
XX WO9217486-A.
PN 15-OCT-1992.
XX 02-APR-1992; 92WO-US02750.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.

PA (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
PI WPI; 1992-366185/44.
DR N-PSDB; AAQ29954.
XX Stimulating proliferation and/or differentiation of primitive
PT mammalian haematopoietic stem cells - using ligand that binds
PT thymidine kinase and flk-1 and flk-2
XX Claim 37; Fig 1a; 94pp; English.
PS The murine foetal liver kinase (flk) -2 sequence was deduced from
XX that of the flk-2 clone obtd. by standard PCR techniques from stem-
CC cell receptor-contg. tissue cDNA libraries. Suitable tissues include
CC foetal liver, spleen or thymus cells or adult marrow or brain cells.
CC The PCR primers used are based on known sections of the flk-2 gene.
CC The murine flk-2 clone may be used in a vector to transform haemato-
CC poietic cells. The thymidine kinase encoded by flk-2 is expressed in
CC primitive but not mature haematopoietic cells. Ligand binding to the
CC TK may be prep'd. which can stimulate proliferation and/or
CC differentiation of primitive haematopoietic cells in vivo. The ligands
CC can stimulate the proliferation of additional primitive stem cells,
CC differentiation into more mature progenitor cells, or both.
CC See also AAR28039-41.
XX
SQ Sequence 992 AA;

Query Match 99.5%; Score 5239; DB 13; Length 992;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 989; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVILETVTNQDLPVKVLSHENNGSSAGKPSRYMRVRS 60
DB 1 MRALQSRDRRLLLVLSVILETVTNQDLPVKVLSHENNGSSAGKPSRYMRVRS 60

QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVWFKHSSLGCCPHFD 120
DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVWFKHSSLGCCPHFD 120

QY 121 LQNRGIVSMALNVTTQAGEYLLHQSERANYTVLFTVNRDTQLYVLRPPYFRKMENQ 180
DB 121 LQNRGIVSMALNVTTQAGEYLLHQSERANYTVLFTVNRDTQLYVLRPPYFRKMENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240

QY 241 ECTKLETDLQAPQSTLPQLFLKVGEPWIRCKAIVHNHGFGLTWELEDKALEEGSYFE 300
DB 241 ECTKLETDLQAPQSTLPQLFLKVGEPWIRCKAIVHNHGFGLTWELEDKALEEGSYFE 300

QY 301 MSTYSNTRMTIRLLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSOEY 360
DB 301 MSTYSNTRMTIRLLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSOEY 360

QY 361 EIDPYEKFCSVRFKAYPRICWIFSQASFCEQRLGLEDGYSLKFCDHKNKPGEIFY 420
DB 361 EIDPYEKFCSVRFKAYPRICWIFSQASFCEQRLGLEDGYSLKFCDHKNKPGEIFY 420

QY 421 AENDDAQFTKMFILNIRKKPOVLANASASQASCSDDGYPLPSWTWKKCDKSPNCTEIP 480
DB 421 AENDDAQFTKMFILNIRKKPOVLANASASQASCSDDGYPLPSWTWKKCDKSPNCTEIP 480

QY 481 EGVNKKANKRVGQWVSSSTLNMSSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
DB 481 EGVNKKANKRVGQWVSSSTLNMSSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540

QY 541 DNISFYATIGLCLPFIWVILVILCHYKKQFRYESQLMIQVTPGLDNEYFYVDFRDEY 600
DB 541 DNISFYATIGLCLPFIWVILVILCHYKKQFRYESQLMIQVTPGLDNEYFYVDFRDEY 600

QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPSENLEFGKVLGSGAFGRVMNATAYGISKTGVSIQVAVKMLKEKADSCKEALM 660

QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKRKFKHRTWTEIFK 720
DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKRKFKHRTWTEIFK 720

QY 721 EHNFSYPTFOAHSSNMPGSRREVQLHPPDLQSLGFGNSIHSSEDEIEYENQRLAESEE 780
DB 721 EHNFSYPTFOAHSSNMPGSRREVQLHPPDLQSLGFGNSIHSSEDEIEYENQRLAESEE 780

QY 781 EDLNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
DB 781 EDLNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840

QY 841 SDSSYVVRGNARLPYKWMAPESLFEIGIYTIKSDVWSYGILLWEISLGVNPGIPVDAN 900
DB 841 SDSSYVVRGNARLPYKWMAPESLFEIGIYTIKSDVWSYGILLWEISLGVNPGIPVDAN 900

QY 901 FYKLQSGFKMEQPPFYATEGIYFVMQSWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960
DB 901 FYKLQSGFKMEQPPFYATEGIYFVMQSWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960

QY 961 SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS 992
DB 961 SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS 992

RESULT 10
AAR81868
ID AAR81868 standard; Protein; 1000 AA.
XX AC AAR81868;
XX DT 29-FEB-1996 (first entry)
XX DE Flk2/flt3 tyrosine kinase receptor.
XX KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;
KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell.
XX OS Mus sp.
XX FH Key
XX FT Domain 1..542
XX FT /label= Extracellular_domain
XX PN W09527062-A1.
XX PD 12-OCT-1995.
XX PF 23-MAR-1995; 95WO-US03718.
XX PR 04-APR-1994; 94US-0222299.
XX PA (GETH) GENENTECH INC.
XX PI Bennett BD, Broz SD, Matthews W, Zeigler FC;
XX N-PSDB; AAT00801.
XX WPI; 1995-358636/46.
XX Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
PT - enhances proliferation of haematopoietic stem cells, in the
PT treatment of hypoplasia, anaemia, etc.
XX Example 1; Page 38-40; 59pp; English.
XX Murine stem cell tyrosine kinase receptor flk2 (also called flt3)
CC (AAR81868) is encoded DNA (AAR00801) obtd. by RT-PCR amplification of
CC RNA isolated from mid-gestation mouse foetal livers. An flk2/flt3
CC extracellular domain-IgG1 Fc fusion protein was used to raise

CC agonist antibodies able to bind to, and activate, flk2/flt3.
CC Such antibodies can enhance repopulation of mature blood cell
CC lineages following chemotherapy, radiotherapy, or bone marrow
CC transplantation.
XX
SQ Sequence 1000 AA;

Query Match 96.9%; Score 5102; DB 16; Length 1000;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRALAQRSDRRLLLLVLSVMILETVTNODLPVVKCVLISHENNGSSACKPSSYRMVRGS 60
DB 1 MRALAQRSDRRLLLLVLSVMILETVTNODLPVVKCVLISHENNGSSACKPSSYRMVRGS 60
QY 61 PEDLOCTPRRSEGVYEAATVEVAESGSLTLQVOLATPGDLSCLWVVKHSSIGCOPHED 120
DB 61 PEDLOCTPRRSEGVYEAATVEVAESGSLTLQVOLATPGDLSCLWVVKHSSIGCOPHED 120
QY 121 LQNRGIVSMAILNVTETQAGEYLLHQISFRANYTVLFTVNVRTDQLYVLRFPYFRKMEQ 180
DB 121 LQNRGIVSMAILNVTETQAGEYLLHQISEAANYTVLFTVNVRTDQLYVLRFPYFRKMEQ 180
QY 181 DALLCISGVPETVWVLCSSHRESCKEEGPAVVRKEEVLHFLFGTD1RCCARNALGR 240
DB 181 DALLCISGVPETVWVLCSSHRESCKEEGPAVVRKEEVLHFLFGTD1RCCARNALGR 240
QY 241 ECTKLFTDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELEDKALEGSYFE 300
DB 241 ESTKLFTDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELEDKALEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFYSSVGRNDTGYYTSCSSKHPQSOSALVTILEKGFINATSSQEEY 360
DB 301 MSTYSTNRTMIRILLAFYSSVGRNDTGYYTSCSSKHPQSOSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYKEFCFSVRKAYPRIRCTWIFSQAPECEQGLDGYSLSKFCDHKNKPGGEYIFY 420
DB 361 EIDPYKEFCFSVRKAYPRIRCTWIFSQAPECEQGLDGYSLSKFCDHKNKPGGEYIFY 420
QY 421 AENDDAQTKFTLNIRKKPOVLANASQASCSGDPPLPSWTWKKCDKSPNCTEIP 480
DB 421 AENDDAQTKFTLNIRKKPOVLANASQASCSGDPPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVNKKANRVFGOWSSSTLNMSEAGKLLVCKAVNSMGTSCETIFLNSPGPFPIQ 540
DB 481 EGVNKKANRVFGOWSSSTLNMSEAGKLLVCKAVNSMGTSCETIFLNSPGPFPIQ 540
QY 541 DNISFYATIGLCLPFIIVLIVLICHKKQFRIESQLOMIQVTPGLDNEYFYVDFRDEY 600
DB 541 DNISFYATIGLCLPFIIVLIVLICHKKQFRIESQLOMIQVTPGLDNEYFYVDFRDEY 600
QY 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAYGISKTGVSIOVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAYGISKTGVSIOVAVKMLKEKADSCKEALM 660
QY 661 SELKMTHLGHHDNIVNLGACTLSGPYLLIFEYCCYGDLLNLRSKRKHRTWTEIFK 720
DB 661 SELKMTHLGHHDNIVNLGACTLSGPYLLIFEYCCYGDLLNLRSKRKHRTWTEIFK 720
QY 721 EHNFSSTPTFAHSNNSMPGSRVOLHPDLQSLGFGNSIHSDEIEYENOKRLAESEE 780
DB 721 EHNFSSTPTFAHSNNSMPGSRVOLHPDLQSLGFGNSIHSDEIEYENOKRLAESEE 780
QY 781 EDLNVLPEDLLCFAYQAKGMEFLFPCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
DB 781 EDLNVLPEDLLCFAYQAKGMEFLFPCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
QY 841 SDSSVVRGNARLPVKWAPESLLEGIYTIKSDVWSYGILLWEIESLGVNYPGIPVDAN 900
DB 841 SDSSVVRGNARLPVKWAPESLLEGIYTIKSDVWSYGILLWEIESLGVNYPGIPVDAN 900
QY 901 FYKLIQSGFKMEQFYATEGYFVVMQSWAFDSRKRPSFPNLTSLGCOLAEAEAMYN 960
DB 901 FYKLIQSGFKMEQFYATEGYFVVMQSWAFDSRKRPSFPNLTSLGCOLAEAEAMYN 960

Db 901 FYKLIQSGFKMEQFYATEGYFVVMQSWAFDSRKRPSFPNLTSLGCOLAEAEAMYN 960
QY 961 -----SIHLPKQAAPQORGLRAQSPQROVKIHRERS 992
DB 961 MCGNVPEHPHSIYONRRPLSREAGS-EPPSPQAQVKIHRERS 1000
RESULT 11
AAR75961
ID AAR75961 standard; Protein; 993 AA.
XX AAR75961;
AC AAR75961;
XX 29-DEC-1995 (first entry)
DT Human STK-1.
DE STK-1; receptor PTK; protein tyrosine kinase.
XX Homo sapiens.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..23
FT /label= signal
FT /note= "unique PTK receptor sequence"
FT Cleavage-site 25..26
FT Domain 35..524
FT /note= "part of extracellular domain contg.
22 Cys residues marked on fig. 1"
FT Modified-site 43..45
FT /label= potential N-linked glycosylation site
FT Modified-site 100..102
FT /label= see above
FT Modified-site 151..153
FT /label= see above
FT Modified-site 250..252
FT /label= see above
FT Modified-site 306..308
FT /label= see above
FT Modified-site 323..325
FT /label= see above
FT Modified-site 351..353
FT /label= see above
FT Modified-site 354..356
FT /label= see above
FT Modified-site 473..475
FT /label= see above
FT Modified-site 502..504
FT /label= see above
FT Region 542..562
FT /label= transmembrane spanning region.
FT Region 708..782
FT /label= kinase insert region
FT Domain 617..622
FT /label= ATP binding domain
FT Domain 835..840
FT /label= WMAPES motifs
FT /note= "cytoplasmic domain"
FT Peptide 808..813
FT /note= "used to design PCR oligos"
FT Peptide 870..875
FT /note= "used to design PCR oligos"
XX WO9519175-A.
PN 20-JUL-1995.
XX 06-JAN-1995; 95WO-US00176.
XX 14-JAN-1994; 94US-0183211.
XX (UJTO) UNIV JOHNS HOPKINS.
XX (UYPE-) UNIV PENNSYLVANIA.
PA

XX Ligand for receptor protein tyrosine kinase - useful for the
PT stimulation of primitive haematopoietic stem cells causing
PT proliferation and/or differentiation
XX
PS Disclosure; Fig 1b; 131pp; English.
XX
CC The sequence corresponds to a human Flk2 (fetal liver kinase)
CC receptor protein-tyrosine-kinase, which is expressed in primitive
CC hematopoietic cells but not in mature hematopoietic cells. The
CC protein is useful in isolation of receptor ligands, which have
CC applications in diagnosis of bone marrow disorders and in
CC stimulating proliferation and/or differentiation of primitive
CC hematopoietic stem cells.
XX
SQ Sequence 993 AA;

Query Match 84.1%; Score 4429.5; DB 16; Length 993;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALAQRSDRLLLVLLVLSVIMLETVTNODLPVKICVLISHENNGSSAGKPSYRWGRGS 60
Db 1 MPALA-RDAGVPLLLVWFSAMIFGTITNODLPVKICVLINHKNDSSVGSYSPNVSSES 59
QY 61 PEDLOCTPROSEGTVYEAATVEVAESGITLOVOLATPGDLSCLWVFKHSSLGCPHF 120
Db 60 PEDLGCLARPOSSGTVPAAAVEVDVSAITLQVLVDAPGNSICLWVFKHSSLNCQPHF 119
QY 121 LQNRGIYSMAILNVTETQAGEYLLHIIQISERANYTVLFTVNVNVDLTQYVLRPFYFKMENQ 180
Db 120 LQNRGVSMVILKMTETOAGEYLLFIQSEATNYTILFTVNSIRNTLLYTLRRPFYFKMENQ 179
QY 181 DALCISEGPEPTVWVLCSSHRESCEGEPNAVYRKEKVLHELFGTDIRCCARNALGR 240
Db 180 DALYCISEGPEPIVWVLCDSQGESCEEPNAVYRKEKVLHELFGTDIRCCARNEILGR 239
QY 241 ECTKLFTIDLQAQOSTLPQLFLKVGPELWIRCKAIHYNHGFGLTWELDEKALEGSEYFE 300
Db 240 ECTRLFTIDLNQTFQTLPQLFLKVGPELWIRCKAVHYNHGFGLTWELDEKALEGSEYFE 299
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVTILEKGFINATSSQBEY 360
Db 300 MSTYSTNRTMIRILLAFVSSVARNDTGYTCTSSSKHPSQSALVTIVGKGFINATSSSEY 359
QY 361 EIDYEFKFCFSVRKAYPRIRCTWIFSOASPCPEORGLDGYISKFCODHKNKPGIYFY 420
Db 360 EIDQYEEFCFSVRKAYPIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHQPGIYFIH 419
QY 421 AENDDAQFTKMTNIRKKPOVLNANASASQSCSDGYPLPSWTWKKSDKSPNCTEIP 480
Db 420 AENDDAQFTKMTNIRKKPOVLNANASQSCSDGYPLPSWTWKKSDKSPNCTEIT 479
QY 481 EGVNKKANRVFGOWVSSSTLNMEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Db 480 EGVNKKANRVFGOWVSSSTLNMEAGKLLVKCCAYNSLGTSCETILLNSPGPFPIQ 539
QY 541 DNISFYATIGLCLPIVVLVILVICHKKQFRYESQLOMIQVTPLDNEYFYVDFRDY 600
Db 540 DNISFYATIGVCLLFIIVVTLVILVICHKKQFRYESQLOMVQVGGSDNEYFYVDFREY 599
QY 601 DLKWEFPRENLEFGKVLGSAFGRVNNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660
Db 600 DLKWEFPRENLEFGKVLGSAFGRVNNATAYGISKTGYSIQVAVKMLKEKADSSEREALM 659
QY 661 SELKMTLGHHDHNVNLLGACTLSGPVYLIFEYCCYGDLLNLRKSKKPHRTWTEIFK 720
Db 660 SELKMTLGLSHENIVNLGACTLSGPVYLIFEYCCYGDLLNLRKSKKPHRTWTEIFK 719
QY 721 EHNFSYPTFOAHSSNMPGSRREVOLHPPDLQSLGFGNSIHSEDEIEYENOKRLAEEREE 780
Db 720 EHNFSYPTFOHPSNMPGSRREVQIHDPDSQISLGHGNSHSEDEIEYENOKRL-EE 777

QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 778 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
QY 841 SDSYVVRGNARLPVKWMAPELSEFEGITYIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
Db 838 SDSNVVVRGNARLPVKWMAPELSEFEGITYIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 897
QY 901 FYKLITQSGFKMQPPYATGEGIVFMQSCWAFDSRKRPSPNLTSLGCCQLAAEAEAC--- 957
Db 898 FYKLITQSGFKMQPPYATGEGIVFMQSCWAFDSRKRPSPNLTSLGCCQLAAEAEAWYQN 957
QY 958 ----TSTSHLPKQAAPQORG-GLRAQSPQROVK 986
Db 958 VDRVSECHTYQNRPPSPREMDLGLLSPQAQVE 991

RESULT 13
AAR67536
ID AAR67536 standard; Protein; 993 AA.
XX AAR67536;
AC AAR67536;
DT 04-JUL-1995 (first entry)
XX Human flk-2.
DE Human flk-2.
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..27
FT Domain 28..543 /label= Sig_peptide
FT Region 544..563 /label= Extracellular_receptor_domain
FT Domain 564..993 /label= Transmembrane_region
FT Domain 994..1000 /label= Intracellular_catalytic_domain
XX US5367057-A.
XX 22-NOV-1994.
XX 02-APR-1991; 91US-0679666.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 19-NOV-1992; 92US-0977451.
XX 30-APR-1993; 93US-0055269.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1995-005894/01.
XX N-PSDB; AAQ79069.
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
PT proliferation and/or stimulation of primitive mammalian
PT haematopoietic stem cells in vitro or in vivo.
XX Disclosure; Fig. 2A-1F; 69pp; English.
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
CC respectively, and the deduced amino acid sequences in AAR67535-37,

CC respectively.

XX Sequence 993 AA;

Query Match 84.1%; Score 4429.5; DB 16; Length 993;

Best Local Similarity 84.1%; Pred. No. 0;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALAQRDRRLLLVLLVSMILETNTDLPVICKVLISHENNSSAGPSSRYMVRGS 60
Db 1 MPALA-RDAGTVPLLVFFSAMFGTTNODLPVICKVLINHNKNDSSVGKSSYPWVSES 59
QY 61 PEDLOCTPRRQSEGTVEAATVEABSGSITLQVLQATPGDLSCLVFKHSSLCGCPHFD 120
Db 60 PEDLGCALRQSSGTVEAATAEVDVSASITLQVLQADPGNISCLVFKHSSLCGCPHFD 119
QY 121 LQNRGIVSMAILNVTQTAGEYLLHQTOSERANTVLTFTVNRVDTQYLVLRPFRKMENQ 180
Db 120 LQNRGVSMVILKMTQTAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMENQ 179
QY 181 DALLCISEGVPETVEMVLCSSHRECKEKGPAVVRKEKVLHLEFGTDIRCCARNALGR 240
Db 180 DALVCISESVPEPIVEMVLCDSOGESCKEESPAVVRKEKVLHLEFGTDIRCCARNELGR 239
QY 241 ECTKLTIDLNAQOSTLPQLFLKVGEPVIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
Db 240 ECTRLFTIDLNOTPQTLFQLKVGEPVIRCKAVHNVHGFGLTWELENKALEEGNYFE 299
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVTILEKGFINATSSOEY 360
Db 300 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVTIVGKGFINATSSSEY 359
QY 361 EIDPYEKFCSVRFKAYPRICRTWISQASFPCEQRLGDEGYSISKFCCHKPKGEYIFY 420
Db 360 EIDQYEEFCVRFKAYPRICRTWISQASFPCEQRLGDEGYSISKFCCHKKHGPGYIFH 419
QY 421 AENDDAQFTKMTFLNTRKRPVLANASASQASCSGYPPLPSWTWKCKSPNCTEEIP 480
Db 420 AENDDAQFTKMTFLNTRRPPVLAELASASQASCSGYPPLPSWTWKCKSPNCTEEIT 479
QY 481 EGVNKKANKRVFGQWVSSTLNMSBAGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
Db 480 EGVNKKANKRVFGQWVSSTLNMSBAGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 539
QY 541 DNISFATITGLCLPFIWLVILCHYKQFRYESQLOMQVTPGLDNEYFYDFRDY 600
Db 540 DNISFATITGVCLLFIWLVILCHYKQFRYESQLOMQVTPGLDNEYFYDFREY 599
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMAATAGISKTGVSIOQAVKMLKEKADSCERL 660
Db 600 DLKWEPPRENLEFGKVLGSGAFGRVMAATAGISKTGVSIOQAVKMLKEKADSERL 659
QY 661 SELKMMTHLGHNDINVLGACTLSPGVLYIFPYCCYGDLLNLYLRSKREKFRHTWTEIFK 720
Db 660 SELKMMTHLGHNDINVLGACTLSPGVLYIFPYCCYGDLLNLYLRSKREKFRHTWTEIFK 719
QY 721 EHNFSYPTFOAHNSMPCSRVQLHPDLQDLSGNGNSIHSDEDEIYENQRLAE 780
Db 720 EHNFSYPTFOAHNSMPCSRVQLHPDLQDLSGNGNSIHSDEDEIYENQRLAE 777
QY 781 EDNLVLTFFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIL 840
Db 778 EDNLVLTFFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIM 837
QY 841 SDSYVVRGNARLPVKWAPESLFEIYTIKSDVSWYSGILLWEIFSGVNPYGPIDAN 900
Db 838 SDSYVVRGNARLPVKWAPESLFEIYTIKSDVSWYSGILLWEIFSGVNPYGPIDAN 897
QY 901 FYKLQSGFKMEOPFVATEIYIVMOSCNWAFDSRKRPSFNLTSFLGCQLAEEAC 957
Db 898 FYKLQSGFKMEOPFVATEIYIVMOSCNWAFDSRKRPSFNLTSFLGCQLAEEAMQN 957
QY 958 ----ITSTHLPKQAAPQRG-GLRAQSPORQVK 986

Db 958 VDCRVSECPHTYQNRPFPSREMDLGLLSPOAQVE 991

RESULT 14

AAR97419

XX ID AAR97419 standard; Protein: 993 AA.

XX AC AAR97419;

XX DT 11-DEC-1996 (first entry)

XX DE Murine foetal liver kinase 2.

XX KW Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;

XX KW monoclonal; antibody; extracellular domain; receptor assay;

XX KW haematopoietic stem cell; ligand; stimulation; proliferation;

XX KW differentiation; treatment; anaemia; bone marrow damage;

XX KW cancer chemotherapy; radiation.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

FT Peptide 1..27

FT Peptide /label= sig_peptide

FT Peptide 28..993

FT Domain /label= mat_peptide

FT Domain 28..544

FT Domain /label= extracellular_domain

FT Domain 545..564

FT Domain /label= transmembrane_domain

FT Domain 565..993

FT Domain /label= intracellular_domain

XX US5548065-A.

XX 20-AUG-1996.

XX 02-APR-1991; 91US-0679666.

XX 19-NOV-1992; 92US-0977451.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 30-APR-1993; 93US-0055269.

XX 31-OCT-1994; 94US-0252517.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1996-392678/39.

XX N-PSDB; AAT38734.

XX Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays.
PT for isolating haematopoietic stem cells expressing receptor and for
PT obtaining ligands

XX Claim 1; Columns 39-48; 50pp; English.

XX The present sequence is murine foetal liver kinase 2 (flk-2),
CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
CC raised against the extracellular portion of flk-2 can be used to
CC assay for flk receptors on the surface of primitive haematopoietic
CC stem cells, and to isolate positive cells. The antibodies can also
CC be used as, or to obtain ligands, which stimulate the proliferation
CC and/or differentiation of stem cells. The ligands can be used, e.g.
CC for treating anaemia, or bone marrow damage resulting from cancer
CC chemotherapy, or radiation.

SQ		Sequence	993 AA;
Query Match		84.1%; Score 4429.5; DB 17; Length 993;	
Best Local Similarity		84.1%; Pred. No. 0;	
Matches 836; Conservative		57; Mismatches 90; Indels 11; Gaps 4;	
Qy	1	MRALAQRSDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSAGKPSRYRMVRGS	60
Db	1	MPALA-RDAGTVPLLVFSAMIFGTITNQDLPVVKCVLINHKNNDSSVGKSSYPMVSES	59
Qy	61	PEDLOCTPQROSEGTVEAATVEAEGSITLQVOLATPGDLSCLWPKHSLGQCPHF	120
Db	60	PEDLGALRPOSSGTVEAAAVEDVSASITLQVLDPAGNISCILWPKHSLGQCPHF	119
Qy	121	LONRGIVSMALNTQAGVLLHIOQSERANYTVLFVNVVDTQLYLRBPYRKMEHQ	180
Db	120	LONRGVSMVLLKMTQAGVLLHIOSEATNYTLFVSVIRNTLLYLRBPYRKMEHQ	179
Qy	181	DALLCISGVPPTVEVWVLCSSHRESCKEKPAAVYKKEKVLHFGTDIRCCARNALGR	240
Db	180	DALVCISESVPEIVWVLCDSQGESCKEESPAVYKKEKVLHFGTDIRCCARNELGR	239
Qy	241	ECTKLFTIDLNAQPOSTLPQLFKVGLPWIRCKAIHVNHGFGLTWELEKALBEGSYFE	300
Db	240	ECTRLFTIDLNOTPTTLPLQLFKVGLPWIRCKAVHVNHGFGLTWELEKALBEGNYFE	299
Qy	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVTILEKGFNATSSQBEY	360
Db	300	MSTYSTNRTMIRILLAFVSSVARNDTGYTCTSSSKHPSQSALVTIVGKGFNATNSSDY	359
Qy	361	EIDPYEFCFSVRKAYPRICRTWIFSAQSPCEORGLEDGYTSKFCGDKHKNKPGYEIFY	420
Db	360	EIDQYEFECFSVRKAYPOICRTWTFSRKSPCEQKGLDNGYSISKFCNHRHQPGYEIFYH	419
Qy	421	AENDDAQFTKMTNIRKKPOVLANASQASCSGSDGYPLPSWTKKCSKSPNCTEIP	480
Db	420	AENDDAQFTKMTNIRKKPOVLAASQASCSGSDGYPLPSWTKKCSKSPNCTEIT	479
Qy	481	EGVWKKANKRVFGOWSSSTFLNSEAGLLVKKCAVNSMGTSCETIFLNSPGPFPIQ	540
Db	480	EGVWNRANKRVFGOWSSSTFLNSEAIGFLVKKCAVNSLGTSCETILLNSPGPFPIQ	539
Qy	541	DNISFYATIGLCLPFIIVVLIIVLICHKKQFRIESQLOMIQVTPLDNEYFYVDPRDYEY	600
Db	540	DNISFYATIGVCLLFIVVLTLLICHKKQFRIESQLOMQVVTGSSDNEYFYVDREYEV	599
Qy	601	DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM	660
Db	600	DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSSEREALM	659
Qy	661	SELKMTHLGHHDNIVNLGACTLSGPVYLIFEXCYCGDLLNLSKREKPHRTWTEIFK	720
Db	660	SELKMTQLGSHENIVNLGACTLSGPVYLIFEXCYCGDLLNLSKREKPHRTWTEIFK	719
Qy	721	EHNFSSTPTQAHNSMPGSRVQLHPDLQSLGFNSGNSHSEDEIEYENQKRLAEDEE	780
Db	720	EHNFSFTPTQSHNSMPGSRVQIHPDSQISLGHNSHSEDEIEYENQKRL-EE	777
Qy	781	EDLNVLTFFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Db	778	EDLNVLTFFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM	837
Qy	841	SDSSVYVRGNARLPVKWMAPELSLEGITIKSDVMSYGILLWEIFSLGVPYGPVVDAN	900
Db	838	SDSNVYVRGNARLPVKWMAPELSLEGITIKSDVMSYGILLWEIFSLGVPYGPVVDAN	897
Qy	901	FYKLIQSCFKMEQPYATGEGYFYVWQSCWAFDSRKRPSPNLTSLGCOLABEAC---	957
Db	898	FYKLIQSCFKMEQPYATGEGYFYVWQSCWAFDSRKRPSPNLTSLGCOLADAEAMYQN	957
Qy	958	---IRTSIHLPLKQAAPOORG-GLRAQSPQROVK 986	
Db	958	VDGRVSECPHYQNRPRPSREMDLGLLSFQAQVE 991	

RESULT 15

AAW19873 standard; Protein; 993 AA.

ID	AAW19873	standard; Protein; 993 AA.
XX	AC	AAW19873;
XX	DT	19-AUG-1997 (first entry)
XX	DE	Human flk-2 receptor.
XX	KW	Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
XX	KW	PTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
XX	KW	haematopoietic hierarchy; extracellular domain; soluble form; ligand;
XX	KW	proliferation; differentiation; mammalian; haematopoietic stem cell;
XX	KW	macrocyclic anaemia; aplastic anaemia; cancer; radiation; mouse.
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	Peptide	1..27
XX	Protein	/note= "Signal peptide"
XX	FT	28..993
XX	FT	/note= "Mature flk-2"
XX	US5621090-A.	
XX	15-APR-1997.	
XX	02-APR-1991;	91US-0679666.
XX	26-JUN-1992;	92US-0906397.
XX	02-APR-1991;	91US-0679666.
XX	28-JUN-1991;	91US-0728913.
XX	15-NOV-1991;	91US-0793065.
XX	24-DEC-1991;	91US-0813593.
XX	(UYPR-) UNIV PRINCETON.	
XX	Lemischka IR;	
XX	WPI: 1997-235228/21.	
XX	N-PSDB: AAT72117.	
XX	Protein containing the extracellular domain of human flk-2 - used	
XX	for identification of primitive haematopoietic cell proliferation	
XX	and differentiation stimulatory ligands, e.g. for treating anaemia	
XX	Claim 1; Fig 1B; 55pp; English.	
XX	This sequence represents human fetal liver kinase 2 (flk2). flk-2 is	
XX	a receptor protein tyrosine kinase (PTK) and is important in transducing	
XX	putative self-renewal signals from the environment. flk-2 is expressed	
XX	in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow.	
XX	and it is thought that flk-2 is expressed in the entire primitive portion	
XX	of the haematopoietic hierarchy. The invention concerns a recombinant	
XX	nucleic acid, preferably mRNA, which encodes a protein containing only	
XX	the extracellular domain of human flk-2 and lacking the flk-2 intra-	
XX	cellular catalytic domain. The resultant protein represents a soluble	
XX	form of flk-2 which is used to isolate specific ligands for flk-2. These	
XX	ligands can be used to stimulate proliferation and/or differentiation of	
XX	mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for	
XX	treatment of macrocytic or aplastic anaemia or bone marrow damage caused	
XX	by cancer treatment or radiation.	
SQ	Sequence	993 AA;

Query Match 84.1%; Score 4429.5; DB 18; Length 993;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

Qy 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSAGKPSRYRMVRGS 60

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	5264	100.0	992	1	US-07-813-593-2		Sequence 2, Appli
2	5264	100.0	992	1	US-07-977-451-2		Sequence 2, Appli
3	5264	100.0	992	1	US-07-946-507-2		Sequence 2, Appli
4	5264	100.0	992	1	US-08-252-517-2		Sequence 2, Appli
5	5264	100.0	992	1	US-07-906-397A-2		Sequence 2, Appli
6	5264	100.0	992	1	US-08-601-891-2		Sequence 2, Appli
7	5264	100.0	992	2	US-09-021-324-2		Sequence 2, Appli
8	5264	100.0	992	5	PCT-US92-02750-2		Sequence 2, Appli
9	5264	100.0	992	5	PCT-US92-05401-2		Sequence 2, Appli
10	5264	100.0	992	5	PCT-US92-09893-2		Sequence 2, Appli
11	5102	96.9	1000	1	US-08-222-299-2		Sequence 2, Appli
12	5102	96.9	1000	2	US-08-434-878-2		Sequence 2, Appli
13	5102	96.9	1000	5	PCT-US95-03718-2		Sequence 2, Appli
14	4436.5	84.3	993	1	US-08-222-299-4		Sequence 4, Appli
15	4436.5	84.3	993	2	US-08-434-878-4		Sequence 4, Appli
16	4436.5	84.3	993	5	PCT-US95-03718-4		Sequence 4, Appli
17	4430.5	84.2	1160	5	PCT-US92-05401-4		Sequence 4, Appli
18	4429.5	84.1	993	1	US-07-977-451-4		Sequence 4, Appli
19	4429.5	84.1	993	1	US-08-252-517-4		Sequence 4, Appli
20	4429.5	84.1	993	1	US-07-906-397A-4		Sequence 4, Appli
21	4429.5	84.1	993	1	US-08-601-891-4		Sequence 4, Appli
22	4429.5	84.1	993	2	US-09-021-324-4		Sequence 4, Appli
23	4429.5	84.1	993	5	PCT-US92-09893-4		Sequence 4, Appli
24	4408.5	83.7	993	1	US-08-183-211-2		Sequence 2, Appli
25	4408.5	83.7	993	5	PCT-US95-00176A-2		Sequence 2, Appli
26	2240.5	42.6	481	4	US-07-912-122-4		Sequence 4, Appli
27	2240.5	42.6	481	5	PCT-US93-06404-4		Sequence 4, Appli

Best Local Similarity 100.0%; Pred. No. 0;		Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MRALQSRDRRLLLLVLSVMIETVTDLPVTKCVLISHENNGSSAGKPSRYRWGRS	60
Db	1	MRALQSRDRRLLLLVLSVMIETVTDLPVTKCVLISHENNGSSAGKPSRYRWGRS	60
Qy	61	PDLOCTPRRQSEGTVYEAATVEAESGSIITQVQLATPGDLSCLVWFVFKHSLGCGPHFD	120
Db	61	PDLOCTPRRQSEGTVYEAATVEAESGSIITQVQLATPGDLSCLVWFVFKHSLGCGPHFD	120
Qy	121	LQNRGIVSMALINVTETQAGEYLLHQSRANYTVLFTVNVNRTDQLYVLRPFYFRKMNQ	180
Db	121	LQNRGIVSMALINVTETQAGEYLLHQSRANYTVLFTVNVNRTDQLYVLRPFYFRKMNQ	180
Qy	181	DALLCISEGVPEPTVWVWVLCSSHRESCKEKGPAVVRKEEKLHFLGTDIRCCARNALGR	240
Db	181	DALLCISEGVPEPTVWVWVLCSSHRESCKEKGPAVVRKEEKLHFLGTDIRCCARNALGR	240
Qy	241	ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIVHNHGFGLTWELDKALEEGSYFE	300
Db	241	ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIVHNHGFGLTWELDKALEEGSYFE	300
Qy	301	MSYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY	360
Db	301	MSYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY	360
Qy	361	EIDPVEKFCFVSFRKAYPIRITWIFQSASFCEQRLGEGYSISKFCDHKNKPGYIFY	420
Db	361	EIDPVEKFCFVSFRKAYPIRITWIFQSASFCEQRLGEGYSISKFCDHKNKPGYIFY	420
Qy	421	ANDDAQFTKMTLNTKRKPOVLANASASQSSDGYPLPSTWTKKCDKSPNCTEEIP	480
Db	421	ANDDAQFTKMTLNTKRKPOVLANASASQSSDGYPLPSTWTKKCDKSPNCTEEIP	480
Qy	481	EGWNNKANRKYFGQWVSSTLNMBSAGKGLLVKCCAYNSMGTSCTEIFLNSPGPPFIQ	540
Db	481	EGWNNKANRKYFGQWVSSTLNMBSAGKGLLVKCCAYNSMGTSCTEIFLNSPGPPFIQ	540
Qy	541	DNISFYATTGLCLPFIWLVILVICHYKKQFRYESQLOMQVTGPLDNEYFYVDFRDY	600
Db	541	DNISFYATTGLCLPFIWLVILVICHYKKQFRYESQLOMQVTGPLDNEYFYVDFRDY	600
Qy	601	DLKWEFPRENLEFGKVLGSGAFGRVWNAAYGISKGTGVSIOQAVKMLKESADCEKALM	660
Db	601	DLKWEFPRENLEFGKVLGSGAFGRVWNAAYGISKGTGVSIOQAVKMLKESADCEKALM	660
Qy	661	SELKMMTHLGHNDINVLNLLGACTLSGPVYLIFEYCCYGDLLNLYLSRKREKPHRTWEIFK	720
Db	661	SELKMMTHLGHNDINVLNLLGACTLSGPVYLIFEYCCYGDLLNLYLSRKREKPHRTWEIFK	720
Qy	721	EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNGNSIHSEDETEYENOKRLAEE	780
Db	721	EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNGNSIHSEDETEYENOKRLAEE	780
Qy	781	EDLNLTFFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Db	781	EDLNLTFFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Qy	841	SDSSVYVRGNARLPVKWMAPESEFEGYTIKSDVWMSYGILLWEIFSLGVNPGIPVDAN	900
Db	841	SDSSVYVRGNARLPVKWMAPESEFEGYTIKSDVWMSYGILLWEIFSLGVNPGIPVDAN	900
Qy	901	FYKLIQSGFKMEQPFYATEGIFYVMSQWAPDSRKRSFNPNTSLFLGCOLAEAEACIRT	960
Db	901	FYKLIQSGFKMEQPFYATEGIFYVMSQWAPDSRKRSFNPNTSLFLGCOLAEAEACIRT	960
Qy	961	SIHLPKQAAPOORGLRAQSPORQVKIHRRS	992
Db	961	SIHLPKQAAPOORGLRAQSPORQVKIHRRS	992

RESULT 2

US-07-977-451-2
; Sequence 2, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-977-451-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALQSRDRRLLLLVLSVMIETVTDLPVTKCVLISHENNGSSAGKPSRYRWGRS 60
Db 1 MRALQSRDRRLLLLVLSVMIETVTDLPVTKCVLISHENNGSSAGKPSRYRWGRS 60

Qy	61	PEDLOCTPRQSGTYYEAAATVEVASGSGTITLQVQATPDGLSCLWVFKHSSLGQCPHFD	120
Db	61	PEDLOCTPRQSGTYYEAAATVEVASGSGTITLQVQATPDGLSCLWVFKHSSLGQCPHFD	120
Qy	121	LQNRGIVSMALLNVTQAGEYLLHQTQSERANVTVLFTVNVRTDQLYLVRRYFRKMNQ	180
Db	121	LQNRGIVSMALLNVTQAGEYLLHQTQSERANVTVLFTVNVRTDQLYLVRRYFRKMNQ	180
Qy	181	DALLCISEGVPEPTBWLVCSSHRECKEKGPAVVRKEEKLHELFGTDIRCCARNALGR	240
Db	181	DALLCISEGVPEPTBWLVCSSHRECKEKGPAVVRKEEKLHELFGTDIRCCARNALGR	240
Qy	241	ECTKLFTIDUNQAPQSTLPOLFUKVGEPLWIRCKAIHVNHGFLTWELDEKDALEBSGYFE	300
Db	241	ECTKLFTIDUNQAPQSTLPOLFUKVGEPLWIRCKAIHVNHGFLTWELDEKDALEBSGYFE	300
Qy	301	MSYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY	360
Db	301	MSYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY	360
Qy	361	EIDPYEKFCFSVRFKAYPRJRCITWIFSQASFCPEQRLGDEGYSISKFDCHKNKPGEYIFY	420
Db	361	EIDPYEKFCFSVRFKAYPRJRCITWIFSQASFCPEQRLGDEGYSISKFDCHKNKPGEYIFY	420
Qy	421	AENDDAQTKMFTLNIRKKPOVLANASASQASCSDDGYPLPSWTMKKCSDKSPNCTEETIP	480
Db	421	AENDDAQTKMFTLNIRKKPOVLANASASQASCSDDGYPLPSWTMKKCSDKSPNCTEETIP	480
Qy	481	EGVWNKKANKRVFGQWVSSSTLNMSBAGKLLVKCCAYNSMGTSCETIFLNSPGPPFFTIQ	540
Db	481	EGVWNKKANKRVFGQWVSSSTLNMSBAGKLLVKCCAYNSMGTSCETIFLNSPGPPFFTIQ	540
Qy	541	DNISFYATIGLCLPFIWVLIVLILCHYKKQFRYESQLOMIQVTPDLNNEYFYVDFRDYEX	600
Db	541	DNISFYATIGLCLPFIWVLIVLILCHYKKQFRYESQLOMIQVTPDLNNEYFYVDFRDYEX	600
Qy	601	DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGTSKGTGVSITQAVKMLKEKADSCKEALM	660
Db	601	DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGTSKGTGVSITQAVKMLKEKADSCKEALM	660
Qy	661	SELKMMTHLGHHDNIYNLGLACTLSGPVYLIFEYCCYCGDLLNLYRSKREKPHRTWTEIFK	720
Db	661	SELKMMTHLGHHDNIYNLGLACTLSGPVYLIFEYCCYCGDLLNLYRSKREKPHRTWTEIFK	720
Qy	721	EHNFSYPTTQAHNSNSMPSGREVQHPHPPDLQLSGFNGNSIHSEDEIYENQKRLAEEBE	780
Db	721	EHNFSYPTTQAHNSNSMPSGREVQHPHPPDLQLSGFNGNSIHSEDEIYENQKRLAEEBE	780
Qy	781	EDLNVLTFEDLLCFAYQVAKGMEFLBFCSCVHRDLAARNVLYTHGKVVKICDFGLARDIL	840
Db	781	EDLNVLTFEDLLCFAYQVAKGMEFLBFCSCVHRDLAARNVLYTHGKVVKICDFGLARDIL	840
Qy	841	SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN	900
Db	841	SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN	900
Qy	901	FYKLIQSGFKMEQFPYATEGIYFVMQSCNAFDSRKRPSPFNLTSLFGCOLAEAEACIRT	960
Db	901	FYKLIQSGFKMEQFPYATEGIYFVMQSCNAFDSRKRPSPFNLTSLFGCOLAEAEACIRT	960
Qy	961	SIHLPKQAAPOQRGGURQAQSPQROVKIHRERS	992
Db	961	SIHLPKQAAPOQRGGURQAQSPQROVKIHRERS	992

RESULT 3

RESOLUTION 3
US-07-946-507-2

US-07-548-307-2
; Sequence 2, Application US/07946507

; Patent No. 5283354

; FACILE NO. 3283334
; GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.

INVENTOR: DEMTSCHUK, THOI K.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

Qy 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLEGGYISIKFCDHKHKNKPGYIFY 420
Db 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLEGGYISIKFCDHKHKNKPGYIFY 420
Qy 421 AENDDAQFTKMTNIRKKPQVLANASASQSCSDGYPPLSWTWKKSCKSPNCTEIP 480
Db 421 AENDDAQFTKMTNIRKKPQVLANASASQSCSDGYPPLSWTWKKSCKSPNCTEIP 480
Qy 481 EGVNKKANRVFGOWVSSSTLNSEACKGLLVKCCAYNSMGTSCETIFLNSPGFPFFIQ 540
Db 481 EGVNKKANRVFGOWVSSSTLNSEACKGLLVKCCAYNSMGTSCETIFLNSPGFPFFIQ 540
Qy 541 DNISFYATIGLCLPFIIVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCLPFIIVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAYGISKTGYISQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAYGISKTGYISQVAVKMLKEKADSCKEALM 660
Qy 661 SELKMTLHGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTETFK 720
Db 661 SELKMTLHGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTETFK 720
Qy 721 EHNFSYPTFOAHNSMPSGREGVQLHPPDLQSLGFGNHSIHSEDEIEYENQKRLAE 780
Db 721 EHNFSYPTFOAHNSMPSGREGVQLHPPDLQSLGFGNHSIHSEDEIEYENQKRLAE 780
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLERKSCVHRDLAARNVLTGHKVYKICDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQVAKGMEFLERKSCVHRDLAARNVLTGHKVYKICDFGLARDIL 840
Qy 841 SDSSVYVRGNARLPVKWMAPESLFEGITIKSDWWSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVYVRGNARLPVKWMAPESLFEGITIKSDWWSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLISQGFKEQFPYATEGYIFVMOŠWAFDSRRKRPFPNLTSLFGCOLAEAEACIRT 960
Db 901 FYKLISQGFKEQFPYATEGYIFVMOŠWAFDSRRKRPFPNLTSLFGCOLAEAEACIRT 960
Qy 961 SIHLPKQAAPQOQGLRAQSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQOQGLRAQSPORQVKIHRERS 992

RESULT 4

US-08-252-517-2
; Sequence 2, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,517
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-252-517-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQSDRRLLLVLSVMILEVTNQDLPVVKVLSHNNSSAGKPSRMYVRGS 60
Db 1 MRALAQSDRRLLLVLSVMILEVTNQDLPVVKVLSHNNSSAGKPSRMYVRGS 60
Qy 61 PEDLOCTPRRQSEGTVYEATVEAESGSITLQVQLATPGDLSCLWYFKHSSLCQPHFD 120
Db 61 PEDLOCTPRRQSEGTVYEATVEAESGSITLQVQLATPGDLSCLWYFKHSSLCQPHFD 120
Qy 121 LQNRGIVSMALLNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPYPKRMENQ 180
Db 121 LQNRGIVSMALLNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPYPKRMENQ 180
Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHLEFCTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHLEFCTDIRCCARNALGR 240
Qy 241 ECTKLFTIDLNAQPOSTLPOLFLKVGEPILWIRKAIHNVHGFGLTWLEBKALEEGSYFE 300
Db 241 ECTKLFTIDLNAQPOSTLPOLFLKVGEPILWIRKAIHNVHGFGLTWLEBKALEEGSYFE 300
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSOEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSOEY 360
Qy 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLEGGYISIKFCDHKHKNKPGYIFY 420
Db 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLEGGYISIKFCDHKHKNKPGYIFY 420
Qy 421 AENDDAQFTKMTNIRKKPQVLANASASQSCSDGYPPLSWTWKKSCKSPNCTEIP 480
Db 421 AENDDAQFTKMTNIRKKPQVLANASASQSCSDGYPPLSWTWKKSCKSPNCTEIP 480

Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEIP 480
Qy 481 EGVNKKANRKYFGQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Db 481 EGVNKKANRKYFGQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Qy 541 DNISFYATIGLCLPPIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 600
Db 541 DNISFYATIGLCLPPIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720
Db 661 SELKMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720
Qy 721 EHNFSYPTFOAHNSNMPGSGREVOLHPPDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780
Db 721 EHNFSYPTFOAHNSNMPGSGREVOLHPPDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780
Qy 781 EDLNVLTPEDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDLNVLTPEDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSVYVRGNARLPVKWAPESLFGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVYVRGNARLPVKWAPESLFGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLIOGFKMEQPYATEGYFVMOQWAFDSRRKPSFPNLTSLFGCOLAEAEACIRT 960
Db 901 FYKLIOGFKMEQPYATEGYFVMOQWAFDSRRKPSFPNLTSLFGCOLAEAEACIRT 960
Qy 961 SIHLPKQAAPQOGLRAQSPQOVKIHRRS 992
Db 961 SIHLPKQAAPQOGLRAQSPQOVKIHRRS 992

RESULT 5

US-07-906-397A-2
; Sequence 2, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906.397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813.593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793.065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728.913
; FILING DATE: 28-JUN-1991

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679.666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28.601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-397A-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Qy 1 MRALAQSRRLRLLLVLSVMILETVTNODLPVVKCVLISHENNGSSAGPSSYRMVGRS 60
Db 1 MRALAQSRRLRLLLVLSVMILETVTNODLPVVKCVLISHENNGSSAGPSSYRMVGRS 60
Qy 61 PEDLQTPRRQSEGTVEATVEAESGSITLQVQLATPGDLSCLVWFKHSSLCGQPHFD 120
Db 61 PEDLQTPRRQSEGTVEATVEAESGSITLQVQLATPGDLSCLVWFKHSSLCGQPHFD 120
Qy 121 LQNRGTVSMAILNVETOAGEYLLHIQSEBANTVLTFTVNRDTQLVLRPFRKMENQ 180
Db 121 LQNRGTVSMAILNVETOAGEYLLHIQSEBANTVLTFTVNRDTQLVLRPFRKMENQ 180
Qy 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEVLHELFGTDIRCCARNALGR 240
Qy 241 ECTKLTIDLNAQOSTLPOLFLKVGEPWIRCKATHVNHGFLGTWLEDKALBEGSYFE 300
Db 241 ECTKLTIDLNAQOSTLPOLFLKVGEPWIRCKATHVNHGFLGTWLEDKALBEGSYFE 300
Qy 301 MSTYSTNRMTIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFNATSSQEEY 360
Db 301 MSTYSTNRMTIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFNATSSQEEY 360
Qy 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFCEQRLGDEGYTSKFCDHKNKPGEIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFCEQRLGDEGYTSKFCDHKNKPGEIFY 420
Qy 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEIP 480
Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEIP 480
Qy 481 EGVNKKANRKYFGQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Db 481 EGVNKKANRKYFGQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Qy 541 DNISFYATIGLCLPPIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 600
Db 541 DNISFYATIGLCLPPIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720
Db 661 SELKMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720
Qy 721 EHNFSYPTFOAHNSNMPGSGREVOLHPPDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780
Db 721 EHNFSYPTFOAHNSNMPGSGREVOLHPPDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780

QY 781 EDNLVLTFFDLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
DB 781 EDNLVLTFFDLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSSVYVRGNARLPVKWMAPESEFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
DB 841 SDSSVYVRGNARLPVKWMAPESEFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEOPFYATGEGIVFMOSCAFDSRKRPSPNLTSFLGCQLAEAEACIRT 960
DB 901 FYKLIOGFKMEOPFYATGEGIVFMOSCAFDSRKRPSPNLTSFLGCQLAEAEACIRT 960
QY 961 SHLPKQAAPOQRGGLRAQSPQVQVXIHRS 992
DB 961 SHLPKQAAPOQRGGLRAQSPQVQVXIHRS 992

RESULT 6

US-08-601-891-2
; Sequence 2, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,891
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-601-891-2
Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALAQRSDRRLLLVVLSVMILETVTNQDLPTVKCVLIHSHENNGSSAGKPSRYMRVGRS 60
DB 1 MRALAQRSDRRLLLVVLSVMILETVTNQDLPTVKCVLIHSHENNGSSAGKPSRYMRVGRS 60
QY 61 PEDLOCTPRQSEGTVYEAAATVEAESGSLTQVQLATPGDLSCLWVFKHSLGCCPHFD 120
DB 61 PEDLOCTPRQSEGTVYEAAATVEAESGSLTQVQLATPGDLSCLWVFKHSLGCCPHFD 120
QY 121 LQNRGIVSMAILNVTQAGEYLLHIQSERANYTVLFTVNVDRDQLYVLRPRYPFKMENQ 180
DB 121 LQNRGIVSMAILNVTQAGEYLLHIQSERANYTVLFTVNVDRDQLYVLRPRYPFKMENQ 180
QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEEVLHLEFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEEVLHLEFGTDIRCCARNALGR 240
QY 241 ECTKLTIDLNQAPQSTLPQLFLKVGPELWIRKAIHNVHGFGLTWELEDKALEGSEYFE 300
DB 241 ECTKLTIDLNQAPQSTLPQLFLKVGPELWIRKAIHNVHGFGLTWELEDKALEGSEYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQBEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQBEY 360
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASPPCEQGLDGYISKFCDHKNKPGEYIFY 420
DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASPPCEQGLDGYISKFCDHKNKPGEYIFY 420
QY 421 AENDDAQFTMFTLNIRKKPQVLANASQASCSGDPYLPFSWTWKKCDSPNCTEIP 480
DB 421 AENDDAQFTMFTLNIRKKPQVLANASQASCSGDPYLPFSWTWKKCDSPNCTEIP 480
QY 481 EGVWNKKANRKFQGWYSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPEFIQ 540
DB 481 EGVWNKKANRKFQGWYSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPEFIQ 540
QY 541 DNISFYATIGLCLPFIIVLVILICHKKQFPRESQLOMIQVTPGLDNEYFYVDFRDEY 600
DB 541 DNISFYATIGLCLPFIIVLVILICHKKQFPRESQLOMIQVTPGLDNEYFYVDFRDEY 600
QY 601 DLKWEFPRENLEFQKVLGSGAFGRVMNATAYGSKTGVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPRENLEFQKVLGSGAFGRVMNATAYGSKTGVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNVLNGLGACTLSGPVYLIFECYCCYDGLLNLYRSKREKFRHTWTFLFK 720
DB 661 SELKMMTHLGHHDNVLNGLGACTLSGPVYLIFECYCCYDGLLNLYRSKREKFRHTWTFLFK 720
QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNNGSIHSEDEIEYENQKRLABEE 780
DB 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNNGSIHSEDEIEYENQKRLABEE 780
QY 781 EDNLVLTFFDLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
DB 781 EDNLVLTFFDLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSSVYVRGNARLPVKWMAPESEFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900

Db 841 SDSSVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPYATGYIVFMOSWAFDSRKRPSFNLTSFLGCOLAAEACIRT 960
Db 901 FYKLIQSGFKMEQPYATGYIVFMOSWAFDSRKRPSFNLTSFLGCOLAAEACIRT 960
Qy 961 SIHLPKQAAPQORGLRAQSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQORGLRAQSPORQVKIHRERS 992

RESULT 7

US-09-021-324-2
Sequence 2, Application US/09021324
Patent No. 5912133
GENERAL INFORMATION:
APPLICANT: Lemishka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-324-2
Query Match 100.0%; Score 5264; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKVLSHNNSSAGKPSRYMVRGS 60
Db 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKVLSHNNSSAGKPSRYMVRGS 60
Qy 61 PEDLOCTPRQSEGTVYEAAATVEAESGSTITLQVQLATPGDLSCLVWFKHSSLGCOPHFD 120
Db 61 PEDLOCTPRQSEGTVYEAAATVEAESGSTITLQVQLATPGDLSCLVWFKHSSLGCOPHFD 120
Qy 121 LQNRGIVSMAILNVTETQAGEVILLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMENQ 180
Db 121 LQNRGIVSMAILNVTETQAGEVILLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMENQ 180
Qy 181 DALLCISGVPPTVWVLCSSHRESCKEEGPAVVRKEEKVHLHELFGTDIRCCARNALGR 240
Db 181 DALLCISGVPPTVWVLCSSHRESCKEEGPAVVRKEEKVHLHELFGTDIRCCARNALGR 240
Qy 241 ECTKLTIDLNOAPOSTLPOLFLKVGEPWIWIRCKATHVNHGGLTWELDKALEEGSYFE 300
Db 241 ECTKLTIDLNOAPOSTLPOLFLKVGEPWIWIRCKATHVNHGGLTWELDKALEEGSYFE 300
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCCSSKHPPOSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCCSSKHPPOSALVTILEKGFINATSSQEEY 360
Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSOASFPCEQRLGLEDGYSISKFDHKNKPGYIFY 420
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSOASFPCEQRLGLEDGYSISKFDHKNKPGYIFY 420
Qy 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGCGYPLPSWTWKKCSKSPNCTEEIP 480
Db 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGCGYPLPSWTWKKCSKSPNCTEEIP 480
Qy 481 EGVWNNKANRKFVGGWVSSSTLNMSBAGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540
Db 481 EGVWNNKANRKFVGGWVSSSTLNMSBAGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540
Qy 541 DNISFYATIGLCPLFIVLVILVILCHIKYKKQFRYESOLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCPLFIVLVILVILCHIKYKKQFRYESOLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660
Qy 661 SELKMWTHLGHHDNIVNLGACTLSPGVLYFEYCCYGLLNLVLRKREKFKHTWTIEFK 720
Db 661 SELKMWTHLGHHDNIVNLGACTLSPGVLYFEYCCYGLLNLVLRKREKFKHTWTIEFK 720
Qy 721 EHNFSYPTFOAHSNNSMPSGREVOLHPDQLSGFNGNSIHSEDEIYEYENQKRLAEEEE 780
Db 721 EHNFSYPTFOAHSNNSMPSGREVOLHPDQLSGFNGNSIHSEDEIYEYENQKRLAEEEE 780
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPYATGYIVFMOSWAFDSRKRPSFNLTSFLGCOLAAEACIRT 960
Db 901 FYKLIQSGFKMEQPYATGYIVFMOSWAFDSRKRPSFNLTSFLGCOLAAEACIRT 960

OY 961 SIHLPKQAAPQORGGLRAQSPOQVQKIHRS 992
Db 961 SIHLPKQAAPQORGGLRAQSPOQVQKIHRS 992
RESULT 8
PCT-US92-02750-2
; Sequence 2, Application PC/TUS9202750
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell
; TITLE OF INVENTION: Receptors And Their Ligands
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: US
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02750
; FILING DATE: 19920402
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FEIT, IRVING N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-02750-2

Query Match 100.0%; Score 5264; DB 5; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MRALAQRDRRLLLVLSVMTLETNTNODLPVICKVLISHENNGSSAGKPPSYRMVRS 60
Db 1 MRALAQRDRRLLLVLSVMTLETNTNODLPVICKVLISHENNGSSAGKPPSYRMVRS 60
OY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLOVQLATPGDLSCLWFKHSSLGCOPHED 120
Db 61 PEDLOCTPRQSEGTVEAATVEAEGSITLOVQLATPGDLSCLWFKHSSLGCOPHED 120
OY 121 LQNRGIVSMAILNVTQAGEYLLHQSERNYVLTFTVNVRTDQLYVLRFPYFRKMENQ 180
Db 121 LQNRGIVSMAILNVTQAGEYLLHQSERNYVLTFTVNVRTDQLYVLRFPYFRKMENQ 180
OY 181 DALLCISEGVPEPTVEMVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVEMVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240
OY 241 ECTKLTIDLNOAPOSTLPOLFLKVCPELWIRCKATHVNHGFLTWELDKALEEGSYFE 300
Db 241 ECTKLTIDLNOAPOSTLPOLFLKVCPELWIRCKATHVNHGFLTWELDKALEEGSYFE 300
OY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKRHPQSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKRHPQSALVTILEKGFINATSSQEEY 360

RESULT 9
PCT-US92-05401-2
; Sequence 2, Application PC/TUS9205401
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05401
; FILING DATE: 19920626
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-05401-2

Query Match		100.0%;	Score 5264;	DB 5;	Length 992;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches		992;	Conservative	0;	Mismatches
				Indels	Gaps
Qy	1	MRALQSRDRRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMVRGS	60		
Db	1	MRALQSRDRRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMVRGS	60		
Qy	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVATPGDLSCLWFKHSSLCQPHFD	120		
Db	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVATPGDLSCLWFKHSSLCQPHFD	120		
Qy	121	LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMEQ	180		
Db	121	LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMEQ	180		
Qy	181	DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEYVLFHFGTDIRCCARNALGR	240		
Db	181	DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEYVLFHFGTDIRCCARNALGR	240		
Qy	241	ECKTLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGTLWEDEKALEEGSYFE	300		
Db	241	ECKTLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGTLWEDEKALEEGSYFE	300		
Qy	301	MSYSTNRTMIRILLAFVSSVGRNDTGYTCSKHPQSALVTILEKGFNATSSQEEY	360		
Db	301	MSYSTNRTMIRILLAFVSSVGRNDTGYTCSKHPQSALVTILEKGFNATSSQEEY	360		
Qy	361	EIDPYEKFCFSVRKAYPRICTWIFSOAQPCEQGLDGYISKFCDHKNKPGEXIFY	420		
Db	361	EIDPYEKFCFSVRKAYPRICTWIFSOAQPCEQGLDGYISKFCDHKNKPGEXIFY	420		
Qy	421	AENDDAQFTKMTNIRKKFOVLNANASQASCSGDPYPLPSWTWKCSKSPNCTEIP	480		
Db	421	AENDDAQFTKMTNIRKKFOVLNANASQASCSGDPYPLPSWTWKCSKSPNCTEIP	480		
Qy	481	EGVWKKANKRVFGQWVSSSTLNNSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ	540		
Db	481	EGVWKKANKRVFGQWVSSSTLNNSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ	540		
Qy	541	DNISFYATIGLCLPFIIVVLIVLICHYKKQPRYESQLQMIQVTPLDNEYFYVDFRDEY	600		
Db	541	DNISFYATIGLCLPFIIVVLIVLICHYKKQPRYESQLQMIQVTPLDNEYFYVDFRDEY	600		
Qy	601	DLKWEFPRENLEFGKVLGSGAFGRMNATAYGSKTGVSIQVAVKMLKEKADSCKEALM	660		
Db	601	DLKWEFPRENLEFGKVLGSGAFGRMNATAYGSKTGVSIQVAVKMLKEKADSCKEALM	660		
Qy	661	SELKMMTHLGHHDNIVNLGACTLSGPVYLIFECYCCGDLNLYRSRKEKPHRTWTEIFK	720		
Db	661	SELKMMTHLGHHDNIVNLGACTLSGPVYLIFECYCCGDLNLYRSRKEKPHRTWTEIFK	720		
Qy	721	EHNFSSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEDEE	780		
Db	721	EHNFSSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEDEE	780		
Qy	781	EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDIL	840		
Db	781	EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDIL	840		
Qy	841	SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN	900		
Db					

Db	841	SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN	900		
Qy	901	FYKLIQSGFKMEQPFYATEGIYFVWQSCWAFDSRKRPSFPNLTSLGCOLABAEACIRT	960		
Db	901	FYKLIQSGFKMEQPFYATEGIYFVWQSCWAFDSRKRPSFPNLTSLGCOLABAEACIRT	960		
Qy	961	SIHLPKQAAPQOGRGLRAQSPQORVKIHRERS	992		
Db	961	SIHLPKQAAPQOGRGLRAQSPQORVKIHRERS	992		

RESULT 10

PCT-US92-09893-2
Sequence 2, Application PC/TUS9209893
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09893
FILING DATE: 19921116
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N. 601
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7PT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09893-2

Query Match		100.0%;	Score 5264;	DB 5;	Length 992;
Best Local Similarity		100.0%;	Pred. No. 0;		
		Matches	992;	Conservative	0;
				Mismatches	0;
				Indels	Gaps
Qy	1	MRALQSRDRRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMVRGS	60		
Db	1	MRALQSRDRRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMVRGS	60		
Qy	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVATPGDLSCLWFKHSSLCQPHFD	120		
Db	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVATPGDLSCLWFKHSSLCQPHFD	120		
Qy	121	LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMEQ	180		
Db	121	LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVNRTQLYVLRPRYFRKMEQ	180		
Qy	181	DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEYVLFHFGTDIRCCARNALGR	240		
Db	181	DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEYVLFHFGTDIRCCARNALGR	240		
Qy	241	ECKTLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGTLWEDEKALEEGSYFE	300		
Db	241	ECKTLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGTLWEDEKALEEGSYFE	300		

```
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOAFSPCEQGLDGYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOAFSPCEQGLDGYISKFCDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSWTWKCDKSPNCTEIEP 480
Db 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSWTWKCDKSPNCTEIEP 480
QY 481 EGWNKANKRVFGQWVSSSTLNMSAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGWNKANKRVFGQWVSSSTLNMSAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCLPFTIVLVILVILCHYKKQFRYESQLOMIQVTPDNEFYVDFRDY 600
Db 541 DNISFYATIGLCLPFTIVLVILVILCHYKKQFRYESQLOMIQVTPDNEFYVDFRDY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWEIFK 720
Db 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWEIFK 720
QY 721 EHNFSYPTFOAHNSMPSGSRREVLPDQLDLSGFNGNSIHSEDEIEYENKRLAEDEE 780
Db 721 EHNFSYPTFOAHNSMPSGSRREVLPDQLDLSGFNGNSIHSEDEIEYENKRLAEDEE 780
QY 781 EDNLVLTEDLLCFAYQAKGMEFLFESKVRDLDLAARNVLVTHGKVKVLCDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQAKGMEFLFESKVRDLDLAARNVLVTHGKVKVLCDFGLARDIL 840
QY 841 SDSSVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLQSGFKMEQPFYATGIEYFVMSQWAFDSRKRPSPNLTSFLGCQLAEEACIRT 960
Db 901 FYKLQSGFKMEQPFYATGIEYFVMSQWAFDSRKRPSPNLTSFLGCQLAEEACIRT 960
QY 961 SHLPKQAPQOQGLRAQSPORQVKIHRERS 992
Db 961 SHLPKQAPQOQGLRAQSPORQVKIHRERS 992

RESULT 11
US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,299
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 879
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-222-299-2
```

```
Query Match 96.9%; Score 5102; DB 1; Length 1000;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
```

```
QY 1 MRALAQSRDRRLLLVLSVMIETVYNQDLPVICKVLISHENNGSSAGKPSRYRMVRS 60
Db 1 MRALAQSRDRRLLLVLSVMIETVYNQDLPVICKVLISHENNGSSAGKPSRYRMVRS 60
QY 61 PEDLOCTPRQSEGTYVEAATVEAESGTTLOVQLATPGDLSCLVWFKHSLSLGCOPHFD 120
Db 61 PEDLOCTPRQSEGTYVEAATVEAESGTTLOVQLATPGDLSCLVWFKHSLSLGCOPHFD 120
QY 121 LQNRGIVSMALNVTTQTAGEYLLHITQSERANTVLTFTVNVRTQLYLVLRPFRKMNQ 180
Db 121 LQNRGIVSMALNVTTQTAGEYLLHITQSERANTVLTFTVNVRTQLYLVLRPFRKMNQ 180
QY 181 DALLCISEGVPEPTVEMVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVEMVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240
QY 241 ECTKLTIDLNAQPOSTLPQLFLKVGEPWIRCAIHVNHGFLGTWELEDKALEEGSYFE 300
Db 241 ECTKLTIDLNAQPOSTLPQLFLKVGEPWIRCAIHVNHGFLGTWELEDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOAFSPCEQGLDGYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOAFSPCEQGLDGYISKFCDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSWTWKCDKSPNCTEIEP 480
Db 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSWTWKCDKSPNCTEIEP 480
QY 481 EGWNKANKRVFGQWVSSSTLNMSAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGWNKANKRVFGQWVSSSTLNMSAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCLPFTIVLVILVILCHYKKQFRYESQLOMIQVTPDNEFYVDFRDY 600
Db 541 DNISFYATIGLCLPFTIVLVILVILCHYKKQFRYESQLOMIQVTPDNEFYVDFRDY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWEIFK 720
Db 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWEIFK 720
```

Qy 721 EHNFSYPTFOAHNSMPSGSRVQLHPPLDQSLGFGNGSIHSEDEIEYENOKRLAESEE 780
Db 721 EHNFSYPTFOAHNSMPSGSRVQLHPPLDQSLGFGNGSIHSEDEIEYENOKRLAESEE 780
Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPELFEIGYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPELFEIGYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Qy 961 -----SIHLPKQAAPQORGLRAOSPOQOVKIHRS 992
Db 961 MGNVPEHPHSIYQNRRLPSREAGS-EPPSPQAQVKIHRS 1000

RESULT 12

US-08-434-878-2

; Sequence 2, Application US/08434878

; Patent No. 5997865

; GENERAL INFORMATION:

; APPLICANT: Bennett, Brian D.

; APPLICANT: Broz, Susan D.

; APPLICANT: Matthews, William

; APPLICANT: Zeigler, Francis C.

; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/434,878

; FILING DATE:

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Hasek, Janet E.

; REGISTRATION NUMBER: 28,616

; REFERENCE/DOCKET NUMBER: 879

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1896

; TELEFAX: 415/952-9881

; TELETYPE: 910/371-7168

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1000 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-434-878-2

Query Match 96.9%; Score 5102; DB 2; Length 1000;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

Qy 1 MRALAQRSDRLLLVLSVMILETVTNODLPVICKVLISHENNGSSAGKPSRYMRVGS 60
|||||

Db 1 MRALAQRSDRLLLVLSVMILETVTNODLPVICKVLISHENNGSSAGKPSRYMRVGS 60
Qy 61 PEDLOCTPRQSEGTVYEAAATVEAESGSITLQVQLATPGDLSCLVWFKHSSJGCGPHFD 120
Db 61 PEDLOCTPRQSEGTVYEAAATVEAESGSITLQVQLATPGDLSCLVWFKHSSJGCGPHFD 120
Qy 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLVLRPFRKMEQ 180
Db 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLVLRPFRKMEQ 180
Qy 181 DALLCISEGVPEPTVWVLCSSHRECKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVWVLCSSHRECKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240
Qy 241 ECTKLFTIDLNOAPOSTLPOLFKVGEPLWIRCKATHVNHGFLTWELDKALEESYFE 300
Db 241 ECTKLFTIDLNOAPOSTLPOLFKVGEPLWIRCKATHVNHGFLTWELDKALEESYFE 300
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSOEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSOEY 360
Qy 361 EIDPYEKFCFSVRFRKAYPRIRCTWIFSQASFCEQRLGLEDGYISKFCCHKNPGEYIF 420
Db 361 EIDPYEKFCFSVRFRKAYPRIRCTWIFSQASFCEQRLGLEDGYISKFCCHKNPGEYIF 420
Qy 421 AENDDAQFTKMETLNIRKKPOVLANASASOASCSOGYPLPSWTWKKCSKSPNCTEIP 480
Db 421 AENDDAQFTKMETLNIRKKPOVLANASASOASCSOGYPLPSWTWKKCSKSPNCTEIP 480
Qy 481 EGWVWKNKANKVFGOWVSSSTLNMSAGGLLVKCCAYNSMGTSCTIFLNSGPPFFIO 540
Db 481 EGWVWKNKANKVFGOWVSSSTLNMSAGGLLVKCCAYNSMGTSCTIFLNSGPPFFIO 540
Qy 541 DNISFYATIGLCLPFIVLVILVICHKKOFRYESOLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCLPFIVLVILVICHKKOFRYESOLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFFRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFFRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFCYCCYGLLNLVLRKREKFRHTWTWEIFK 720
Db 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFCYCCYGLLNLVLRKREKFRHTWTWEIFK 720
Qy 721 EHNFSYPTFOAHNSMPSGSRVQLHPPLDQSLGFGNGSIHSEDEIEYENOKRLAESEE 780
Db 721 EHNFSYPTFOAHNSMPSGSRVQLHPPLDQSLGFGNGSIHSEDEIEYENOKRLAESEE 780
Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPELFEIGYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPELFEIGYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Qy 961 -----SIHLPKQAAPQORGLRAOSPOQOVKIHRS 992
Db 961 MGNVPEHPHSIYQNRRLPSREAGS-EPPSPQAQVKIHRS 1000

RESULT 13

PCT-US95-03718-2

; Sequence 2, Application PC/TUS9503718

; GENERAL INFORMATION:

; APPLICANT: GENENTECH, INC.

; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03718
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 879PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1000 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-03718-2

Query Match 96.9%; Score 5102; DB 5; Length 1000;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
QY 1 MRALAQRSDRLLLLVLSVMILETVTQDLPVVKCVLISHENNGSSAGKPSRYMRVRS 60
DB 1 MRALAQRSDRLLLLVLSVMILETVTQDLPVVKCVLISHENNGSSAGKPSRYMRVRS 60
QY 61 PEDLOCTPRQSEGVYEAATVEAESGSIITLOVQLATPGDLSCLVVFKHSSIGCOPHFD 120
DB 61 PEDLOCTPRQSEGVYEAATVEAESGSIITLOVQLATPGDLSCLVVFKHSSIGCOPHFD 120
QY 121 LQNRGIVSMALNVETQAGEYLLHTQSPRANTVLTFTVNRDQTLVLRRLRPYFRKMEHQ 180
DB 121 LQNRGIVSMALNVETQAGEYLLHTQSPRANTVLTFTVNRDQTLVLRRLRPYFRKMEHQ 180
QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEEGPAVVRKEKVLHFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEEGPAVVRKEKVLHFGTDIRCCARNALGR 240
QY 241 ECTKLFTIDLQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300
DB 241 ECTKLFTIDLQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFCPEQRLGDEGYSISKFDHKNKPGYIFY 420
DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFCPEQRLGDEGYSISKFDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTFLNIRKKPQVLANASQAQSCSDGYPLPSTWTKKCDKSPNCTEEP 480
DB 421 AENDDAQFTKMTFLNIRKKPQVLANASQAQSCSDGYPLPSTWTKKCDKSPNCTEEP 480
QY 481 EGVWNNKANRKFVGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540

DB 481 EGVWNNKANRKFVGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540
QY 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKQPRYEQSOLQMIQVTGPDLDNEFYFYVDFRDY 600
DB 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKQPRYEQSOLQMIQVTGPDLDNEFYFYVDFRDY 600
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLSKRKEFHRTWTEIFK 720
DB 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLSKRKEFHRTWTEIFK 720
QY 721 EHNFSYPTFOAHSSNMPGSRVOLHPDLQSGFNGNSIHSEDEIEYENOKRLAESEE 780
DB 721 EHNFSYPTFOAHSSNMPGSRVOLHPDLQSGFNGNSIHSEDEIEYENOKRLAESEE 780
QY 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
DB 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
QY 841 SDSYVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
DB 841 SDSYVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
QY 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSPNLTSLGCLAEAEACIRT 960
DB 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSPNLTSLGCLAEAEACIRT 960
QY 961 -----SIHLPKQAAPQGGRLAQQSPQROVKIHRERS 992
DB 961 MCGNVPEHPSIYQNRRLPSREAGS-EPPSPQAQVKIHRERS 1000

RESULT 14
US-08-222-299-4
Sequence 4, Application US/08222299
Patent No. 5635388
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Broz, Susan D.
APPLICANT: Matthews, William
APPLICANT: Zeigler, Francis C.
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,299
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-299-4

Query Match 84.3%; Score 4436.5; DB 1; Length 993;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

Qy 1 MRALQSRDRRLLLVLSVMIETVTTNODLPVVKVLSHENNGSSAGKPSRYMRVRS 60
Db 1 MPALA-RDGGQLPLLVWFSAIFGTITNODLPVVKVLSHENNGSSAGKPSRYMRVRS 59

Qy 61 PEDLOCTPRROSEGVYEAATVEVAESGSIITQVOLATPGDLSCLVWFKHSLGCOQPHD 120
Db 60 PEDLGCLARPOSSGTVYEAATVEVAESGSIITQVOLATPGDLSCLVWFKHSLGCOQPHD 119

Qy 121 LQNRGIVSMALNTVETQAGEYLLHIQSERANYTVLTVNVDRDTQYVLRPRYPRKMEQ 180
Db 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYITLFTVSTINTLLYLRPRYPRKMEQ 179

Qy 181 DALLCISEGVPEPTVEVWVLCSSHRESCKEESGPAVVKKEKVLHFGTDIRCCARNALGR 240
Db 180 DALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGMDIRCCARNELGR 239

Qy 241 ECTKLFIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEBSGYFE 300
Db 240 ECTKLFIDLNQOTPTLPQLFKVGEPLWIRCKAVHNVHGFGLTWELENKALEBSGYFE 299

Qy 301 MSTYSTNRTMIRILAFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFNATSSQEX 360
Db 300 MSTYSTNRTMIRILAFVSSVGRNDGYTCTSSSKHPSQSALVTIVKGFNATSSSEY 359

Qy 361 EIDPYEKFCFSVREKAYPRICRTWIFSOAFPCQEGLEDGYSISKFDHKNKPGCEYIF 420
Db 360 EIDQYEEFCFSVREKAYPIQIRCTWTFESKSPCEQKGLDNGYSISKFCNKHQPGCEYIF 419

Qy 421 AENDDAQFTKFTNIRKKPOVLNANASQASCSGDPPLPSWTKKCDKSPNCTEIP 480
Db 420 AENDDAQFTKFTNIRKKPOVLNANASQASCSGDPPLPSWTKKCDKSPNCTEIT 479

Qy 481 EGVNKKANRVFGOWSSSTPLNSEACKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Db 480 EGVNKKANRVFGOWSSSTPLNSEACKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 539

Qy 541 DNISFYATIGLCLPFIYVVLIVLICHYKKQFRESQLOMIQVTPGLDNEYFYVDFRDY 600
Db 540 DNISFYATIGVCLLFIYVVLIVLICHYKKQFRESQLOMIQVTPGLDNEYFYVDFREY 599

Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATYGISKTGVSIOVAVKMLKEKADSCKEALM 660
Db 600 DLKWEFPRENLEFGKVLGSGAFGRVMNATYGISKTGVSIOVAVKMLKEKADSSEREALM 659

Qy 661 SELKMTHLGHHDINVLGACTLSGPVYLIFECYCGDILNLYRSKREKFRHTWTEIFK 720
Db 660 SELKMTQLGSHENINVLGACTLSGPVYLIFECYCGDILNLYRSKREKFRHTWTEIFK 719

Qy 721 EHNFSYPTFOAHNSMMPGSRVQLHPPLDQLSGFNNSIHSDEIEYENOKRLAEDEE 780
Db 720 EHNFSYPTFOAHNSMMPGSRVQLHPPLDQLSGFNNSIHSDEIEYENOKRLAEDEE 777

Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 778 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837

Qy 841 SDSYVVRGNARLPVKMAPESLFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
Db 838 SDSYVVRGNARLPVKMAPESLFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 897

Qy 901 FYKLIOGSKMEOPFYATEGYFYVWQSCWAFDSRKRPSFNLTSFLGCOLADEAEACIRT 960
Db 900 FYKLIOGSKMEOPFYATEGYFYVWQSCWAFDSRKRPSFNLTSFLGCOLADEAEACIRT 956

Db 898 FYKLIQNGFKMDQPEYATEEIIYIMQSCWAFDSRKRPSFNLTSFLGCOLADEAEAMYQ- 956

Qy 961 STHLPKQAP-----QRRGGLRAQSPQOVK 986

Db 957 NVDPGVSECPHTYQNRPRFSREMDLGLLSPQAQVE 991

RESULT 15
US-08-434-878-4
Sequence 4, Application US/08434878
Patent No. 5997865
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Broz, Susan D.
APPLICANT: Matthews, William
APPLICANT: Zeigler, Francis C.
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,878

FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-434-878-4

Query Match 84.3%; Score 4436.5; DB 2; Length 993;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

Qy 1 MRALQSRDRRLLLVLSVMIETVTTNODLPVVKVLSHENNGSSAGKPSRYMRVRS 60
Db 1 MPALA-RDGGQLPLLVWFSAIFGTITNODLPVVKVLSHENNGSSAGKPSRYMRVRS 59

Qy 61 PEDLOCTPRROSEGVYEAATVEVAESGSIITQVOLATPGDLSCLVWFKHSLGCOQPHD 120
Db 60 PEDLGCLARPOSSGTVYEAATVEVAESGSIITQVOLATPGDLSCLVWFKHSLGCOQPHD 119

Qy 121 LQNRGIVSMALNTVETQAGEYLLHIQSERANYTVLTVNVDRDTQYVLRPRYPRKMEQ 180
Db 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYITLFTVSTINTLLYLRPRYPRKMEQ 179

Qy 181 DALLCISEGVPEPTVEVWVLCSSHRESCKEESGPAVVKKEKVLHFGTDIRCCARNALGR 240
Db 180 DALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGMDIRCCARNELGR 239

Qy 241 ECTKLFIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEBSGYFE 300
Db 240 ECTKLFIDLNQOTPTLPQLFKVGEPLWIRCKAVHNVHGFGLTWELENKALEBSGYFE 299

Db 240 ECTRLFTIDLNTQTTLPLQLKAVGEPLTRCAVHVNHGFGLTWELENKALEEGNYFE 299
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
Db 300 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTIVEKGFINATSSSEY 359
Qy 361 EIDPYEKFCFSVRKAYPRICRTWIFSOASPPCQORGLDGYISKFCDHKNKPGXEYIF 420
Db 360 EIDQYEEFCFSVRKAYPOICRTWTFSRKSPCQKGLDNGYSISKFCNHHKHQGEYIFH 419
Qy 421 AENDDAQFTKMTLNIRKKPOVLNANASASQSCSDGYPLPSWTWKKCSDKSPNCTEIP 480
Db 420 AENDDAQFTKMTLNIRKKPOVLNANASASQSCSDGYPLPSWTWKKCSDKSPNCTEIT 479
Qy 481 EGVNKKANRVFGQWSSVTLNMSKAGLLVKKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Db 480 EGVNKKANRVFGQWSSVTLNMSKAGLLVKKCCAYNSMGTSCETIFLNSPGPFPIQ 539
Qy 541 DNISFYATIGLCLPFIVLVILVICHKKYKOFRYESOLQMIQVTPLDNEYFYVDFRDY 600
Db 540 DNISFYATIGLCLPFIVLVILVICHKKYKOFRYESOLQMIQVTPLDNEYFYVDFRDY 599
Qy 601 DLKWEFFPRENLEFGKVLGSGAGFGRVMNATAYGISKTVYSIQVAVKMLKEKADSCKEALM 660
Db 600 DLKWEFFPRENLEFGKVLGSGAGFGRVMNATAYGISKTVYSIQVAVKMLKEKADSCSEREALM 659
Qy 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720
Db 660 SELKMTQLGSHENIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK 719
Qy 721 EHNFSYPTFOAHNSMPSGREVQLHPDQLGCFNGNSTHSEDEIEYENOKRLAEDEE 780
Db 720 EHNFSYPTFOAHNSMPSGREVQLHPDQLGCFNGNSTHSEDEIEYENOKRLAEDEE 777
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840
Db 778 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIM 837
Qy 841 SDSSYVVRGNARLPVKWAPESLFEIGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
Db 838 SDSSYVVRGNARLPVKWAPESLFEIGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 897
Qy 901 FYKLQSGFKMEQPYATEGIIYVMSQWAFDSRKRSPNLTSLGCOLAEAEACIET 960
Db 898 FYKLQSGFKMEQPYATEGIIYVMSQWAFDSRKRSPNLTSLGCOLAEAEACIET 956
Qy 961 STHLPKQAP-----QORGLRAQSPQOVK 986
Db 957 NVDGPVSECPHTYQNRPFRENDLGLLSPOAQVE 991

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:34:40 ; Search time 24.4877 Seconds
(without alignments)
4017.505 Million cell updates/sec

Title: US-09-919-408-2
Perfect score: 5264
Sequence: 1 MRALAQRSDRLLLVLSV.....RGGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues
Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5264	100.0	992	10	US-09-919-408-2
2	5264	100.0	992	10	US-09-872-136-2
3	4429.5	84.1	993	10	US-09-919-408-4
4	4429.5	84.1	993	10	US-09-872-136-4
5	1223	23.2	975	9	US-10-192-867-2
6	1221.5	23.2	972	10	US-09-944-807-10
7	1196	22.7	976	9	US-10-099-895-32
8	1196	22.7	976	9	US-10-132-867-4
9	1152.5	21.9	1088	9	US-09-961-403-4
10	1152.5	21.9	1089	10	US-09-769-987-2
11	1152.5	21.9	1089	10	US-09-919-497-90
12	1152.5	21.9	1089	10	US-09-866-510-2
13	1150.5	21.9	1089	9	US-09-955-363-36
14	1149.5	21.8	1089	10	US-09-866-510-10
15	1148.5	21.8	1089	10	US-09-866-510-4
16	1147.5	21.8	1089	10	US-09-866-510-8
17	1146.5	21.8	1089	10	US-09-866-510-6
18	1078	20.5	1090	10	US-09-866-510-14
19	1078	20.5	1106	9	US-09-955-363-2

20	1078	20.5	1106	10	US-09-866-510-22	Sequence 22, Appl
21	1074	20.4	1106	10	US-09-866-510-16	Sequence 16, Appl
22	1073	20.4	1106	10	US-09-866-510-20	Sequence 20, Appl
23	1072	20.4	1106	10	US-09-866-510-18	Sequence 18, Appl
24	998.5	19.0	1338	9	US-10-059-585-44	Sequence 44, Appl
25	961.5	18.3	1356	9	US-10-022-939-2	Sequence 2, Appl
26	961.5	18.3	1356	9	US-10-100-405A-2	Sequence 2, Appl
27	960.5	18.2	1356	9	US-09-969-037-7	Sequence 7, Appl
28	954.5	18.1	1298	10	US-09-982-610-33	Sequence 33, Appl
29	954.5	18.1	1363	9	US-09-375-248-2	Sequence 2, Appl
30	952.5	18.1	1368	9	US-10-105-901-34	Sequence 34, Appl
31	952	18.1	1363	9	US-09-375-248-19	Sequence 19, Appl
32	943.5	17.9	1362	9	US-10-105-901-33	Sequence 33, Appl
33	941	17.9	386	9	US-09-939-833-6	Sequence 6, Appl
34	941	17.9	386	10	US-09-939-754-6	Sequence 6, Appl
35	941	17.9	386	10	US-09-939-832-6	Sequence 6, Appl
36	937.5	17.8	1367	10	US-09-766-678-2	Sequence 2, Appl
37	936.5	17.8	1363	9	US-10-105-901-32	Sequence 32, Appl
38	926.5	17.6	1367	10	US-09-919-408-6	Sequence 6, Appl
39	926.5	17.6	1367	10	US-09-872-136-6	Sequence 6, Appl
40	862	16.4	367	9	US-09-939-833-9	Sequence 9, Appl
41	862	16.4	367	10	US-09-939-754-9	Sequence 9, Appl
42	862	16.4	367	10	US-09-939-832-9	Sequence 9, Appl
43	856	16.3	367	9	US-09-939-833-12	Sequence 12, Appl
44	856	16.3	367	10	US-09-939-754-12	Sequence 12, Appl
45	856	16.3	367	10	US-09-939-832-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-919-408-2
; Sequence 2, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0., Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-919-408-2

Query Match 100.0%; Score 5264; DB 10; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRDRLLLVLSVMILETVTNDLPVIRKVLISHENNGSSACKPSSRYMRVGS 60
Db 1 MRALAQRDRLLLVLSVMILETVTNDLPVIRKVLISHENNGSSACKPSSRYMRVGS 60

Qy 61 PEDLOCTPRQSEGTVEAATVEAEGSITLOVOLATPGDLSCLVFKHSSLGCOHPED 120
Db 61 PEDLOCTPRQSEGTVEAATVEAEGSITLOVOLATPGDLSCLVFKHSSLGCOHPED 120

Qy 121 LQNRGIVSMAILNVETQAGEYLLHIOSEKRVNTVLFTVNVNRTOLYVLRPPYFRKMNQ 180
Db 121 LQNRGIVSMAILNVETQAGEYLLHIOSEKRVNTVLFTVNVNRTOLYVLRPPYFRKMNQ 180

Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEEGPAVVRKEEKVLHFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVWVLCSSHRESCKEEGPAVVRKEEKVLHFGTDIRCCARNALGR 240

Qy 241 ECTKLTFTDLNOAPOSTLPOLFKVGEPLWIRCKAIHVNHGFLTWELDKALEGSYFE 300
Db 241 ECTKLTFTDLNOAPOSTLPOLFKVGEPLWIRCKAIHVNHGFLTWELDKALEGSYFE 300

Qy 301 MTSYSTNRTMIRILLAFVSSVGRNDTGYYTCSSKHPSOSALVTILEKGFINATSSOEY 360
Db 301 MTSYSTNRTMIRILLAFVSSVGRNDTGYYTCSSKHPSOSALVTILEKGFINATSSOEY 360

Qy 361 EIDPYEKFCSVRFKAYPRICTWIFSQAFPCQORLEDGYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKFCSVRFKAYPRICTWIFSQAFPCQORLEDGYISKFCDHKNKPGYIFY 420

Qy 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSTWTKKCDKSPNCTEIP 480
Db 421 AENDDAQFTKMTLNIRKKPOVLANASASQSCSDGYPLPSTWTKKCDKSPNCTEIP 480

Qy 481 EGVNKNKARKVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGVNKNKARKVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540

Qy 541 DNISFYATIGLCLPFTVILVILICHYKKOFRYESOLQIQTGTPLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCLPFTVILVILICHYKKOFRYESOLQIQTGTPLDNEYFYVDFRDY 600

Qy 601 DLKWEPPRENLEFGKVLGSAFGVRVNNATAYIGSKTGVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEPPRENLEFGKVLGSAFGVRVNNATAYIGSKTGVSIQVAVKMLKEKADSCKEALM 660

Qy 661 SELKMTHTLGHHDNIYNILGACTLSGPVYLIFEYCCYGDLLNLYRSKREKHFHTWEIFK 720
Db 661 SELKMTHTLGHHDNIYNILGACTLSGPVYLIFEYCCYGDLLNLYRSKREKHFHTWEIFK 720

Qy 721 EHNFSYPPTFOAHSSNMPGSRVQLHPPLDQLSGFNGNSIHSEDETEYENQKRLAE 780
Db 721 EHNFSYPPTFOAHSSNMPGSRVQLHPPLDQLSGFNGNSIHSEDETEYENQKRLAE 780

Qy 781 EDNLNLTFFEDLLCFAYQVAKGMEFLEKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 781 EDNLNLTFFEDLLCFAYQVAKGMEFLEKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840

Qy 841 SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVWYSGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVWYSGILLWEIFSLGVNYPGIPVDAN 900

Qy 901 FYKLIQSGFKMEOPFYATEGIYVWQSCNAFDSRKPSFNLTSFLGCCQLAEAEACIRT 960
Db 901 FYKLIQSGFKMEOPFYATEGIYVWQSCNAFDSRKPSFNLTSFLGCCQLAEAEACIRT 960

Qy 961 SIHLPKQAAPQORGLRAQSPORQVRIHRERS 992
Db 961 SIHLPKQAAPQORGLRAQSPORQVRIHRERS 992

RESULT 2
US-09-872-136-2
; Sequence 2, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872.136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208.786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021.324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977.451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906.397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813.593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793.065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728.913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679.666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-872-136-2

Query Match 100.0%; Score 5264; DB 10; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRDRLLLLVLSVLMLETVTQDLPVVKCVLISHENNGSSAGKPSRYRMVGRS 60
Db 1 MRALAQRDRLLLLVLSVLMLETVTQDLPVVKCVLISHENNGSSAGKPSRYRMVGRS 60

Qy 61 PEDLOCTPRROSEGTVYEAATVEAEGSITLQVOLATPGDLSCLLVKFKHSLGCOQPHD 120
Db 61 PEDLOCTPRROSEGTVYEAATVEAEGSITLQVOLATPGDLSCLLVKFKHSLGCOQPHD 120

Qy 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVNVDLTOLYVLRPFYFRKMENQ 180
Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVNVDLTOLYVLRPFYFRKMENQ 180

Qy 181 DALLCISEGVEPTVWVLCSSHRESCKEKGPAVYRKEEVLHFLGCTDIRCCARNALGR 240
Db 181 DALLCISEGVEPTVWVLCSSHRESCKEKGPAVYRKEEVLHFLGCTDIRCCARNALGR 240

Qy 241 ECTKLFTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEGSGYFE 300
Db 241 ECTKLFTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEGSGYFE 300

Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPQSALVTLERGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPQSALVTLERGFINATSSQEEY 360

Qy 361 EIDPYEKEFCFVRKAYPRICTWIFSOASPCOEGLDGYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKEFCFVRKAYPRICTWIFSOASPCOEGLDGYISKFCDHKNKPGYIFY 420

Qy 421 AENDDAQOFTKFTLNIRKKPOVLANASQASCSDDGYPLPSWTWKKCKSDKSPNTEETIP 480
Db 421 AENDDAQOFTKFTLNIRKKPOVLANASQASCSDDGYPLPSWTWKKCKSDKSPNTEETIP 480

Qy 481 EGVNKKANRVKFGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Db 481 EGVNKKANRVKFGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540

Qy 541 DNISFYATIGLCLPIVVLVILICHYKQFYESOLOMIQVTGPDNDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCLPIVVLVILICHYKQFYESOLOMIQVTGPDNDNEYFYVDFRDY 600

Qy 601 DLKWEFPRENLEFGKVGSGAFGRVNNATAYGISKTGYISQVAVKMLKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVGSGAFGRVNNATAYGISKTGYISQVAVKMLKADSCKEALM 660

Qy 661 SELKMTHLGHHDNIVNLIGACTLSGPVYLPEYCCYGDLLNLYRSKREKFRHTWTETPK 720
Db 661 SELKMTHLGHHDNIVNLIGACTLSGPVYLPEYCCYGDLLNLYRSKREKFRHTWTETPK 720

Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPPDLQSLGFGNSIHSEDEIEYENOKRLAEDEE 780
Db 721 EHNFSYPTFOAHNSMPSGSEVQLHPPDLQSLGFGNSIHSEDEIEYENOKRLAEDEE 780

Qy 781 EDNLVLTEDLLCFAYQVAKGHEFLFEKSCVHRDLAARNVLVTHGKVKYKICDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQVAKGHEFLFEKSCVHRDLAARNVLVTHGKVKYKICDFGLARDIL 840

Qy 841 SDSSVVRGNARLPVKWAPESLFGITIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWAPESLFGITIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900

Qy 901 FYKLIQSGFKMEQPFYATEGIYFVWQSWAFDSRKRPSFNLTSFLGCCLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATEGIYFVWQSWAFDSRKRPSFNLTSFLGCCLAEAEACIRT 960

Qy 961 SIHLPKQAAPQORGGGLRAQSPQOVKIHRRS 992
Db 961 SIHLPKQAAPQORGGGLRAQSPQOVKIHRRS 992

RESULT 3
US-09-919-408-4
; Sequence 4, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Query Match 84.1%; Score 4429.5; DB 10; Length 993;
Best Local Similarity 84.1%; Pred. No. 2.3e-297;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALQSRDRRLLLVLSVMIETVNDQLPVTKVLSHENNGSSAGKPSRYMRVRS 60
DB 1 MPALA-RDAGTVPLLVVFSAMIFGITNODLPVTKVLSHENNGSSAGKPSRYMRVRS 59
QY 61 PEDLOCTPRQSGTVEAATVEAEGSITLOVQLATPGDLSCLWVFKHSSLCQPHFD 120
DB 60 PEDLGALRQSSGTVEAATVEAEGSITLOVQLATPGDLSCLWVFKHSSLCQPHFD 119
QY 121 LQNRGIYSMAILNVTQAGEVLLHLSQSRANYTVLFTVNVDRDTOLYVLRPRYKRMENQ 180
DB 120 LQNRGIYSMAILNVTQAGEVLLHLSQSRANYTVLFTVNVDRDTOLYVLRPRYKRMENQ 179
QY 181 DALLCISEGVPEPTVEWLVCSHRECKEESPAVYKKEKVLHLEFGTDIRCCARNALGR 240
DB 180 DALVCISEVPEPTVEWLVCSHRECKEESPAVYKKEKVLHLEFGTDIRCCARNALGR 239
QY 241 ECTKLTIDNQAQSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300
DB 240 ECTRLFTIDNQAQSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 299
QY 301 MSTYSTNRTMIRLLAFVSSVGRNDGYTCSSSKHPSOSALVTILEKGFINATSSOEY 360
DB 300 MSTYSTNRTMIRLLAFVSSVGRNDGYTCSSSKHPSOSALVTILEKGFINATSSOEY 359
QY 361 EIDPYEKFCSVRFKAYPRIRCTWIFSOASFPCEQRLGEGYSISKFCOHKKNKPGYIEY 420
DB 360 EIDPYEKFCSVRFKAYPRIRCTWIFSOASFPCEQRLGEGYSISKFCOHKKNKPGYIEY 419
QY 421 AENDDAQFTKMTLNTRKPKQVLANASASQASCSGYPPLPSWTWKCSKSPNCTEETP 480
DB 420 AENDDAQFTKMTLNTRKPKQVLANASASQASCSGYPPLPSWTWKCSKSPNCTEETP 479
QY 481 EGVNKKANKRVFGQWSSSTLNMSAGLLVKKCCAYNSMTGSCETIFLNSGPPPEIQ 540
DB 480 EGVNKKANKRVFGQWSSSTLNMSAGLLVKKCCAYNSMTGSCETIFLNSGPPPEIQ 539
QY 541 DNISFATIGLCLPFTVILVILCHYKQFRYESQLOMQIVTQPDNNEYFYDFRDY 600
DB 540 DNISFATIGLCLPFTVILVILCHYKQFRYESQLOMQIVTQPDNNEYFYDFRDY 599
QY 601 DLKWEPPRENLEFGKVLGSGAGRVNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
DB 600 DLKWEPPRENLEFGKVLGSGAGRVNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 659
QY 661 SELKMTWHLGHNDINVLGACLTSPVLYLFEYCCYGDLLNLYRSKREKPHFTWTEIFK 720
DB 660 SELKMTWHLGHNDINVLGACLTSPVLYLFEYCCYGDLLNLYRSKREKPHFTWTEIFK 719
QY 721 EHNFSYPTFOAHNSMPSREVQLPDLQSLGFGNGSIHSEDEIYENQKRLAEDEE 780
DB 720 EHNFSYPTFOAHNSMPSREVQLPDLQSLGFGNGSIHSEDEIYENQKRLAEDEE 777
QY 781 EDNLVLTFFDLCLFAYQVAKGMEFEFKSCVHRDLAARNVLTGKVVKICDFGLARDIL 840
DB 780 EDNLVLTFFDLCLFAYQVAKGMEFEFKSCVHRDLAARNVLTGKVVKICDFGLARDIL 837
QY 841 SDSYVVRGNARLPVKWMAPESEFEGITTKSDVWSYIGILLWELFSLGVNPPYGPIDAN 900
DB 840 SDSYVVRGNARLPVKWMAPESEFEGITTKSDVWSYIGILLWELFSLGVNPPYGPIDAN 897
QY 901 FYKLQSGFKMEOPFVATEGIIYVMSQWAFDSKRPSPNLTSLGCOLAEAEAC--- 957
DB 900 FYKLQSGFKMEOPFVATEGIIYVMSQWAFDSKRPSPNLTSLGCOLAEAEAC--- 957
QY 958 ----IRTSIHLPKQAAPQORG-GIRAQSPQOVK 986
DB 958 VDRGVSECPHYQNRNRFREMDLGLLSPQAQVE 991

RESULT 4

US-09-872-136-4

; Sequence 4, Application US/09872136

; Patent No. US20020119545A1

GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-872-136-4

Query Match 84.1%; Score 4429.5; DB 10; Length 993;

Best Local Similarity 84.1%; Pred. No. 2,3e-297;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALQSRDRRLLLVLSVMIETVNDQLPVTKVLSHENNGSSAGKPSRYMRVRS 60

DB 1 MPALA-RDAGTVPLLVVFSAMIFGITNODLPVTKVLSHENNGSSAGKPSRYMRVRS 59

QY 61 PEDLOCTPRQSGTVEAATVEAEGSITLOVQLATPGDLSCLWVFKHSSLCQPHFD 120

DB 60 PEDLGALRQSSGTVEAATVEAEGSITLOVQLATPGDLSCLWVFKHSSLCQPHFD 119

QY 121 LQNRGIYSMAILNVTQAGEVLLHLSQSRANYTVLFTVNVDRDTOLYVLRPRYKRMENQ 180

Db 166 PKAGIMIKSVKRAYHRLCLHCSVDQEG-----KSVLSEKFIKVRPAFK-----A 210
QY 245 LFTIDLNAQPOSTLPOLFLKVGEPDLWIRCKAIHVNHGFLTWELEDK--ALEE----- 295
Db 211 VPVSVSKA-----SYLLREGGEFTVTCTIKDVSSSVYSTWKRENSQTKLOEKYNSWHH 264
QY 296 GSYFEMSTYSTNRTMIRILLAFVSSVGRNDGTGYTSCSSKHPQSALVTILE--KGFIN 352
Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCVYANNTFGSANVTTLVVDKGFIN 314
QY 353 A-TSSQEEYIDPYEKFCFSVRFKAYPR-IRCTWIFSQASF--PCEQRLGLEDGYISKFC 408
Db 315 IFPMINTTVFVNDGENDVLIYEYEAFFKPEHQOQWIMYNNRTFTDKWEDYPKSENESIRYV 374
QY 409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKPKQVLA--NASASQASCSDDGYP 459
Db 375 SELHLTRKGTGGTYTFLVNSDVNAIAAFNVYNTKPEILTYDRLVNGMLQCAAGFP 434
QY 460 LPSWTWKKCKDKSPNCTEEI-PEGVWNNKANRKFVGOWSSSTLNMSSEAGKGLLVKCCAY 518
Db 435 EPTIDWYFCGTEQRCASVLPVDVQTLNSSGPPFGKLVQSSIDSSAFKNGVTECKAY 494
QY 519 NSMGTSCETIFLNSPGPPFIQDN-----ISFYATIGLCLPFIIVLVILIC 564
Db 495 NDVGKT--SAVEN---FAFGKNKEQIHPHTLFTPLLLIGFVIVAGM---MCIIVMILT 544
QY 565 HKYKKQFRIESQLOMI--QVTGPLDNEYFYVDFRDYEYDLKWEFPRENLEFGKVLGSGAFG 623
Db 545 YKYLQKPMYEQWKVVEEING---NNVYIDPTQLPDYDHKWEFPNRLSFGKTLGAGAFG 601
QY 624 RVMNATAYGISTGVSIQAVKMLKEKADSCKEALMSKMMTHLGHHDNIVNLGACT 683
Db 602 KVEATAYGLKSDAAMTVAVKMLKPSAHLTEREALMSKLVLSYLGNNHMINVLGACT 661
QY 684 LSGPVYLIFCYCCYGDLLNLRKREKFRHTWTEIFKEHNFSSYPTFQAHSNNSMPGSR 743
Db 662 IGGPTLVITEYCCYGDLLNLRKRRDSFICSQEDHAEAAALYKNLLHSSKSSCSDSTNEY 721
QY 744 VOLHPPDLQOL---SGFNNGSIHSEDEIYEYENOKRLAEEDLNLVTFEDLLCFAYQVAK 800
Db 722 MDMPGVSVVYVPTKADRRSVRIGSYIERDVTPTAIMEDDELALDL---EDLLSFSYQVAK 778
QY 801 GMEFLEFKSVHRDLAARNLVTHGKVKYKICDGLARDILSDSSVYVVRGNARLPVKWAP 860
Db 779 GMAFLASKNCHIRDLAARNILLTHGRITKICDGLARDIKNDSNYYVVGKARLPVKWAP 838
QY 861 ESLFEGYITIKSDVWSYGILLWEIFSLGVNPPYGPIDVANDFYKLIQSGFKMEQPFYATEG 920
Db 839 ESIFNCVYTFESDWSYGIWELFSLGSSPYGMPVDSKFKYKMKEGFRMLSPHAPAE 898
QY 921 IYFVWQSCWAFDSKRSPFNLTSLGQCLAE 953
Db 899 MYDIMKTCWDADPLKRPTRFQIIVOLIEKOISES 931

RESULT 8

US-10-192-867-4
; Sequence 4, Application US/10192867
; Publication No. US2003008466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192.867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

; LENGTH: 976
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-192-867-4

Query Match 22.7%, Score 1196; DB 9; Length 976;
Best Local Similarity 31.4%; Pred. No. 3.le-74;
Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;

QY 36 CVLISHENNGSAGKPSRYRMVGRSPEDLQCTPRQSECTVYEAATVEVAESGITLQVO 95
Db 12 CVLLLLLRVQGTSSQPS-----VSPGPPSPSIHPGKSDILVRVGD-EIRLL 57
QY 96 LATPCDLSCLVFKHSSLCQPHFDLQNRGIVSMALLNVTETQAGEYLLHHIOSEIRANYTV 155
Db 58 CTDPGFVK--WTFE---ILDETENKQNEWITEKA-----EATNTKYTCTNKHGJNSIY 108
QY 156 LFTVNRD--TOLYVLRPRYRKMEKMODALLCI-----SEG-----VPE 192
Db 109 VF---VRDPAKFLVDRSLYKEDNDTLVRCPLTDPVNTYSLKGCQKPLPKDLRFIPD 165
QY 193 PTVKVLCSHRE-----SCKEEGPAVVRKEBKVLHFLGTDIRCCARNALGRECTK 244
Db 166 PKAGIMIKSVKRAYHRLCLHCSVDQEG-----KSVLSEKFIKVRPAFK-----A 210
QY 245 LFTIDLNAQPOSTLPOLFLKVGEPDLWIRCKAIHVNHGFLTWELEDK--ALEE----- 295
Db 211 VPVSVSKA-----SYLLREGGEFTVTCTIKDVSSSVYSTWKRENSQTKLOEKYNSWHH 264
QY 296 GSYFEMSTYSTNRTMIRILLAFVSSVGRNDGTGYTSCSSKHPQSALVTILE--KGFIN 352
Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCVYANNTFGSANVTTLVVDKGFIN 314
QY 353 A-TSSQEEYIDPYEKFCFSVRFKAYPR-IRCTWIFSQASF--PCEQRLGLEDGYISKFC 408
Db 315 IFPMINTTVFVNDGENDVLIYEYEAFFKPEHQOQWIMYNNRTFTDKWEDYPKSENESIRYV 374
QY 409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKPKQVLA--NASASQASCSDDGYP 459
Db 375 SELHLTRKGTGGTYTFLVNSDVNAIAAFNVYNTKPEILTYDRLVNGMLQCAAGFP 434
QY 460 LPSWTWKKCKDKSPNCTEEI-PEGVWNNKANRKFVGOWSSSTLNMSSEAGKGLLVKCCAY 518
Db 435 EPTIDWYFCGTEQRCASVLPVDVQTLNSSGPPFGKLVQSSIDSSAFKNGVTECKAY 494
QY 519 NSMGTSCETIFLNSPGPPFIQDN-----ISFYATIGLCLPFIIVLVILIC 564
Db 495 NDVGKT--SAVEN---FAFGKNKEQIHPHTLFTPLLLIGFVIVAGM---MCIIVMILT 544
QY 565 HKYKKQFRIESQLOMI--QVTGPLDNEYFYVDFRDYEYDLKWEFPRENLEFGKVLGSGAFG 623
Db 545 YKYLQKPMYEQWKVVEEING---NNVYIDPTQLPDYDHKWEFPNRLSFGKTLGAGAFG 601
QY 624 RVMNATAYGISTGVSIQAVKMLKEKADSCKEALMSKMMTHLGHHDNIVNLGACT 683
Db 602 KVEATAYGLKSDAAMTVAVKMLKPSAHLTEREALMSKLVLSYLGNNHMINVLGACT 661
QY 684 LSGPVYLIFCYCCYGDLLNLRKREKFRHTWTEIFKEHNFSSYPTFQAHSNNSMPGSR 743
Db 662 IGGPTLVITEYCCYGDLLNLRKRRDSFICSQEDHAEAAALYKNLLHSSKSSCSDSTNEY 721
QY 744 VOLHPPDLQOL---SGFNNGSIHSEDEIYEYENOKRLAEEDLNLVTFEDLLCFAYQVAK 800
Db 722 MDMPGVSVVYVPTKADRRSVRIGSYIERDVTPTAIMEDDELALDL---EDLLSFSYQVAK 778
QY 801 GMEFLEFKSVHRDLAARNLVTHGKVKYKICDGLARDILSDSSVYVVRGNARLPVKWAP 860
Db 779 GMAFLASKNCHIRDLAARNILLTHGRITKICDGLARDIKNDSNYYVVGKARLPVKWAP 838
QY 861 ESLFEGYITIKSDVWSYGILLWEIFSLGVNPPYGPIDVANDFYKLIQSGFKMEQPFYATEG 920
Db 839 ESIFNCVYTFESDWSYGIWELFSLGSSPYGMPVDSKFKYKMKEGFRMLSPHAPAE 898
QY 921 IYFVWQSCWAFDSKRSPFNLTSLGQCLAE 953
Db 899 MYDIMKTCWDADPLKRPTRFQIIVOLIEKOISES 931


```
Db 137 LVIVEDDDSAIIPCRITDPTETVTLHNSGVVPA SYDSRQGFNGTFTVGPYICEATVKGK 196
Qy 241 ECTKLTIDLNOAPOSTLPOFL-----KVGEPLWIRCKAIHVNHGFLTW-----E 287
Db 197 ---KFQTFPNVYALKATSELDLEALKTVYKSGTIVVTC-AVFNNEVDLOWTYPGE 252
Qy 288 LEDKALEGYSFEMSTYTNRTMIR-----ILLAFVSSVGR---NDGYTCSSSK--- 335
Db 253 VKGKI-----TMLEIKVPSIKLVITVPEATVKDSDGYECAARQATR 297
Qy 336 --HPSQSALVTILEKGF--INATSSOEYEIDPYEKFCFSVRKAYPRIRCTWIFSOASF 391
Db 298 EVKEMKKVTISVHEKGFIEIKPTFSOLE-AVNLHEVHFVVEVRAYPPRISMLKNNLTL 356
Qy 392 PCEQRLGEGYSISKFC-----HKNK-----PGEYIFYAENDDAQTKMFTLN 435
Db 357 -----IENLTETDVEKIQEIRYRSKLIRAKEEDSGHYTIVAQNEDAVKSYTFELL 410
Qy 436 IRKKPQVL-----ANASASQASCSGDYPLPSWTWKCKDPSNCTEIEPEGVWNKKA 488
Db 411 TQVPSSILDLVDDHGGTGVRCVTAEGTPLDIEWMICKD-IKKCNNETS---WTILA 466
Qy 489 NRKVFQWVSS-----STLNMSSEAGKGLLVKCCAYNSMGTSCETIFLNSPG 534
Db 467 NN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAKNLGAENRELKIVA-- 518
Qy 535 PPFPTQDNISFYATIGLCLPFIVLVILVICHYKQFYESOLQMQVTPDNEFYFVD 594
Db 519 --PTLRSELTVAAAVLVIVIIISLIVLVHWKQPRVEIRWRVIESISPDGHEIYVD 576
Qy 595 FRDYEDLKWEPFRENLEFGVLGSGAGFGRVMNATAYGKTSKGVSIQVAVKMLKEKADSC 654
Db 577 PMQLPYDSRWEPFRDGLVLRVLSGAGFVKVVEGTAYGLSRSPQVMKVAVKMLKPTARSS 636
Qy 655 EKEALMSELKMTHLGHNDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSRKEFP--- 711
Db 637 EKQALMSELKIMTHLGPLHLNIVNLLGACTKSGPIIYITEYCFYGDVLYNLHKNRDSFLSH 696
Qy 712 -----HRTWTEIFEKHNFSYPTFOAHNSMPP--GSEVOLHPP 749
Db 697 HPEKPKKELDIFGLNPADESTRYSYVILSFENNGDYMDMQADTTQVPMLEKREYSKID 756
Qy 750 LDQLSGFNGNSIHSEDEIYENOKRLAEEDDLNLTFEDLLCFAYQVAKGMEFELEFS 809
Db 757 IORSLYDRPASVKKKSMDSSEVKNLLSDNSEGLTLL---DLLSFTYQVARGMEFLASKN 813
Qy 810 CVHRDLAARNVLVTHGKVKVICKDFGLARDILSDSSVYVRGNARLPVKWMAPESLFEGIYT 869
Db 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNVVSGSTFLPVKWMAPESIFDNLTY 873
Qy 870 IKSDVWSYGILLWEIFSLGVPNPGIPVDANFYKLIQSGFKMEQPFYATGYTFVMOQSCW 929
Db 874 TSLDWSYGILLWEIFSLGVPNPGIPVDANFYKLIQSGFKMEQPFYATGYTFVMOQSCW 933
Qy 930 AFDSKRPSFNLTSFL 946
Db 934 NSEPEKRPFSYHLSEIV 950
```

RESULT 11

```
US-09-919-497-90
; Sequence 90, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; CURRENT APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-90

Query Match 21.9%; Score 1152.5; DB 10; Length 1089;
Best Local Similarity 32.1%; Pred. No. 3.6e-71;
Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

Qy 94 VOLATPGDLSCLWFKHSSLGCPHDLQNRGIVSMAILNVNTEQAGEY--LLHIQSER 151
Db 39 VOLNSFSLR-----FGESEVSMQ--YPMSEESSDVEIRN-EENSGSLFTVLEVSASA 92
Qy 152 NTVLFTVNRDTQLVLRPRYFRKMNODALLCISEG-----VPEPTV-----EW 197
Db 93 AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIVPDPVAFVPLGMDY 136
Qy 198 VLCSHRES---CK---EEGPAVVRKEEVL-----HELFGTDIRCCARNALGR 240
Db 137 LVIVEDDDSAIIPCRITDPTETVTLHNSGVVPA SYDSRQGFNGTFTVGPYICEATVKGK 196
Qy 241 ECTKLTIDLNOAPOSTLPOFL-----KVGEPLWIRCKAIHVNHGFLTW-----E 287
Db 197 ---KFQTFPNVYALKATSELDLEALKTVYKSGTIVVTC-AVFNNEVDLOWTYPGE 252
Qy 288 LEDKALEGYSFEMSTYTNRTMIR-----ILLAFVSSVGR---NDGYTCSSSK--- 335
Db 253 VKGKI-----TMLEIKVPSIKLVITVPEATVKDSDGYECAARQATR 297
Qy 336 --HPSQSALVTILEKGF--INATSSOEYEIDPYEKFCFSVRKAYPRIRCTWIFSOASF 391
Db 298 EVKEMKKVTISVHEKGFIEIKPTFSOLE-AVNLHEVHFVVEVRAYPPRISMLKNNLTL 356
Qy 392 PCEQRLGEGYSISKFC-----HKNK-----PGEYIFYAENDDAQTKMFTLN 435
Db 357 -----IENLTETDVEKIQEIRYRSKLIRAKEEDSGHYTIVAQNEDAVKSYTFELL 410
Qy 436 IRKKPQVL-----ANASASQASCSGDYPLPSWTWKCKDPSNCTEIEPEGVWNKKA 488
Db 411 TQVPSSILDLVDDHGGTGVRCVTAEGTPLDIEWMICKD-IKKCNNETS---WTILA 466
Qy 489 NRKVFQWVSS-----STLNMSSEAGKGLLVKCCAYNSMGTSCETIFLNSPG 534
Db 467 NN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAKNLGAENRELKIVA-- 518
Qy 535 PPFPTQDNISFYATIGLCLPFIVLVILVICHYKQFYESOLQMQVTPDNEFYFVD 594
Db 519 --PTLRSELTVAAAVLVIVIIISLIVLVHWKQPRVEIRWRVIESISPDGHEIYVD 576
Qy 595 FRDYEDLKWEPFRENLEFGVLGSGAGFGRVMNATAYGKTSKGVSIQVAVKMLKEKADSC 654
Db 577 PMQLPYDSRWEPFRDGLVLRVLSGAGFVKVVEGTAYGLSRSPQVMKVAVKMLKPTARSS 636
Qy 655 EKEALMSELKMTHLGHNDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSRKEFP--- 711
Db 637 EKQALMSELKIMTHLGPLHLNIVNLLGACTKSGPIIYITEYCFYGDVLYNLHKNRDSFLSH 696
Qy 712 -----HRTWTEIFEKHNFSYPTFOAHNSMPP--GSEVOLHPP 749
Db 697 HPEKPKKELDIFGLNPADESTRYSYVILSFENNGDYMDMQADTTQVPMLEKREYSKID 756
Qy 750 LDQLSGFNGNSIHSEDEIYENOKRLAEEDDLNLTFEDLLCFAYQVAKGMEFELEFS 809
Db 757 IORSLYDRPASVKKKSMDSSEVKNLLSDNSEGLTLL---DLLSFTYQVARGMEFLASKN 813
Qy 810 CVHRDLAARNVLVTHGKVKVICKDFGLARDILSDSSVYVRGNARLPVKWMAPESLFEGIYT 869
Db 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNVVSGSTFLPVKWMAPESIFDNLTY 873
Qy 870 IKSDVWSYGILLWEIFSLGVPNPGIPVDANFYKLIQSGFKMEQPFYATGYTFVMOQSCW 929
Db 874 TSLDWSYGILLWEIFSLGVPNPGIPVDANFYKLIQSGFKMEQPFYATGYTFVMOQSCW 933
Qy 930 AFDSKRPSFNLTSFL 946
Db 934 NSEPEKRPFSYHLSEIV 950
```



```

; LENGTH: 1089 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-955-363-36

```

Query Match	21.98;	Score 1150.5;	DB 9;	Length 1089;
Best Local Similarity	32.29;	Pred. No. 5e-71;		
Matches 313;	Conservative 154;	Mismatches 327;	Indels 177;	Gaps 32;
QY	94	VQATPGDLSCLWFKHSSIGCQCFHDLQNGIYSMAILNWTETQAGEY--LLHIQSERA	151	
DB	39	VQJNSSFSLR-----FGSEVSWQ--YPSBEESDVEIRN--EENSGLGFTVLEVSSASA	92	
QY	152	NYTVLFTVNVWRDTQLYVLRRPYFRKMENQDALLCISEG-----VPEPTV-----EW	197	
DB	93	AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIYVDPDVAFVPLGMDY	136	
QY	198	VLCSSHRES-----CK--EEGPAVVRKEKVL-----HELFGTDICCCARNALGR	240	
DB	137	LVIIVEDDDSAIIPCRTPDPETPVILHNSGVVPASYSRQGFENGFTTVPYICEATVKGK	196	
QY	241	ECKLFTIDLNAQPOSTLPQLFL-----KVGEPLWIRCKATHVNHGFLTW----E	287	
DB	197	---KFQTIPEFNVAALKATSELDLEMEALKTVYKSGETIWTCT-AVFNNEVVDLQWTPGE	252	
QY	288	LEDKALEEGSYFENSTYSTNTMIRILLAFVSSVGR---NDTYGTCSSK-----HPSQ	339	
DB	253	VKGKGI-----TILEETKPSIKLVITLTVPETVKDSGDGYECARQATREVKEMK	303	
QY	340	SALVTILEKGF--INATSSQBEYETDPYEKFCFSYRFRKAYPRIRCTWIFSOASFPCEQRG	397	
DB	304	KVTISVHEKGFGIEIKTFPSQLE-AVNLHEVHKVFVVEVRAYPPIRISLWLNKLTL-----	356	
QY	398	LEDGYSTSKPCD-----HKMK-----PGEIYIYAENDDAQFTKMTLNRKKPQ	441	
DB	357	IENUTEITDDVEKIQEIRYRSKLLIRAKEEDSGHYITVAQNEDAVKSYYTTELLTQVPSS	416	
QY	442	VL-----ANASASQACSSGDYPLPSWTWKKCDKSPNCTEETPEGVWKNKANKRVFG	494	
DB	417	ILDVLDDHHGSTGQTVRCATAGTPLPDIEWKCD-IKKCNNETS---WILANN-----	468	
QY	495	QWVSS-----STLNNSSEAGKLLLVKCCAYNSMGTSETIFLNSPGPFPIQ	540	
DB	469	--VSNIIITEHSRDRSTVEGRVTFKAVEETIAVRCLAKNLLGAENREKLVA---PTLR	522	
QY	541	DNISFYATIGLCLFFIVVLIVLICHKYKQFYRESOLOMIQVTPGLDNEYFVDFRDYEX	600	
DB	523	SELTVAAVLLVIVIVISLIVLVIMQKPRYELRWRVIESISPDGHEIYIVDPMQLPY	582	
QY	601	DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGKISGTGVSIQVAVKMLKEKADSCKEALM	660	
DB	583	DSRWEFFRDGLVGLRVLGSGAFKVGVECTAYGLSRQSPVMKVAVKMLKPTARSSKEQALM	642	
QY	661	SELKMPHGLHHDNVLNLLGACTLSGPVYLFFCYCCYGDLLNLYLRSKEKF-----	711	
DB	643	SELKIMTHLGLHMLNVLNLLGACTSGPIIYITEYCFYGDVLNLYLHKNRDSFLSHHPEKPK	702	
QY	712	-----HRTWTEIFKEHNFSYPTFOAHNSMSP--GSRVEQLHPLPDQLSG	755	
DB	703	KELDIFGLNPADESTRSVILSFENNGYDMKQADTQYVPMLEKVEYSKYSDIORSLY	762	
QY	756	FNGNSIHSDEIEIYENOKRLAEEREEEDLNVITFEDLLCFAYOVAKGMFEFLFKSCVHRDL	815	
DB	763	DRPASYKKKSLDSEVKNLLSDNNSGUTLL--DLSFTYQVARGMEFLASKNCVHRDL	819	
QY	816	AARNVLVTHGWKVICDFGLARDIILSDSSYVVRGNARLPVKWMAPESEFEGITYTIKSDVW	875	
DB	820	AARNVLAQGIKVICDFGLARDIIMHDSNYSYKSGTFLPVKWMAPESIFDNLITYTSLDVW	879	
QY	876	SYGILLWEIFSLGYNPPIGVADNANFYKLIQSGFKMEQPFYATEGIVFVMSQWAFDRSK	935	

```

Db      880  SYGILLWEIFSLGCTPYPGMMVDSTFYNKISGYRMAKPDHATSEVYEIMVKWNSPEK 939
Qy      936  RPSFPNLTSL 946
      |||| :|| :
Db      940  RPSFVHLSEIV 950

RESULT 14
US-09-866-510-10
; Sequence 10, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISE
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-10

```

Query Match	21.8%;	Score 1149.5;	DB 10;	Length 1089;
Best Local Similarity	32.0%;	Pred. No. 5.8e-71;		
Matches 313;	Conservative 155;	Mismatches 320;	Indels 189;	Gaps 33;
Qy	94	VOLATPGDLSCLWFKHSSLCQPHFDLQNRGIYSMAILNVTETOAGBY--LLHQISERA	151	
Db	39	VOLNSFSRLC---FGESEVSWQ--YPMSEESDVEIRN--EENNSGLFVTVLEVSSASA	92	
Qy	152	NYTVLFTVNVNRTQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW	197	
Db	93	AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIYVPPDVAVFPLGMTDY	136	
Qy	198	VLCSSHRES----CK---EEGPVAVRKEEVL-----HELFGTDIRCCARNALGR	240	
Db	137	LVIVEDDDSAIIPCKRTTDPETPVTVLHNSGVVPASYDSROGFNGTFTVGPYICATVKKG	196	
Qy	241	ECTKLFTIDLNOAQPOSTLPQLFL-----KVGEPLWIRCKAIHNHNGFLW----E	287	
Db	197	---RFQTIFFNVYALKATSELDEMEALKTVYKGEIYVTC-AVFNEVVDLQWTPGE	252	
Qy	288	LEDKALEEGSFEMSTYSTNTMTIR-----ILAFVSSVGR-----NTGYTTCSSSK---	335	
Db	253	VKGKI-----TMLIEIKVPISIKVLTYLTPTAEATVKDSGDGYECAARQATR	297	
Qy	336	--HFSOSALVTLEKGF--INATSSOEYEIDPYEKKCFSVRFKRAYPIRCTWLFWSQASF	391	
Db	298	EVKEMKKVTIYSHKGFIEIKPTFSQLE-AVNLHEVKHFVVVEVRAYPPPIRISWLNKNNLTL	356	
Qy	392	PCOEGLGEDGYSISKFCD-----HKNK-----PGEYIEAENDDAQFTKMETLN	435	
Db	357	-----IENLTIETDVEKIQEIRYSRKLIRAKEEDSGHYTIVAQNEDAVKSYTFELL	410	
Qy	436	IRKQPVL-----ANASASQACSSDGYPLPSWTWKKCKSDKSPNCTEETIPEGVWNKKA	488	
Db	411	TQVPSIILDLVDHGHSTGGTGVTRCTAEGTLPDIEMWICKD-IKKCNNETS---WTILA	466	
Qy	489	NRKVFGQWVSS-----STLNMSEACKGLLVKVCAYNSMGTSCTETIFLNSPG	534	
Db	467	NN-----VSNLIIEHRSRSTVEGRVTFKAVEETAVRCLANKLGAENREKLVA--	518	
Qy	535	PFPTIQDNISFYATIGLCLPFIVVLIVLICHKKYKQRFYESQLOMQVTQPLDNEYFYVD	594	
Db	519	--PTLRSBLTVAAGVLLVTVIISLIVLVVIMKQRYEIRWRVIESISPDGHEIYVD	576	

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:27:39 ; Search time 26.4867 Seconds
(without alignments)
3600.506 Million cell updates/sec

Title: US-09-919-408-2
Perfect score: 5264
Sequence: 1 MRALAQRSDRRLLLVLSV.....RGGLRAQSPQROVKIHRRS 992

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	A39931	protein-tyrosine k
2	5102	96.9	1000	S18827	Flt3 protein - mou
3	4408.5	83.7	993	A36873	protein-tyrosine k
4	1258.5	23.9	980	TVCTMD	macrophage colony-
5	1247	23.7	941	TVNVMQ	protein-tyrosine k
6	1232.5	23.4	978	A49814	protein-tyrosine k
7	1223	23.2	975	TVMSKT	macrophage colony-
8	1221.5	23.2	972	TVHUMD	macrophage colony-
9	1220	23.2	954	I51703	c-kit-related kina
10	1215	23.1	977	I45877	protein-tyrosine k
11	1210	23.0	978	S16385	macrophage colony-
12	1198.5	22.8	976	TVMSMD	macrophage colony-
13	1196	22.7	976	TVHUKT	macrophage colony-
14	1188.5	22.6	975	J30816	macrophage colony-
15	1181	22.4	960	JN0677	protein-tyrosine k
16	1157.5	22.0	1088	PPRTGA	platelet-derived g
17	1152.5	21.9	1089	PFHUGA	platelet-derived g
18	1132	21.5	1087	I51552	platelet-derived g
19	1123.5	21.3	1089	S33727	platelet-derived g
20	1098	20.9	1089	PFMSRB	platelet-derived g
21	1078	20.5	1106	PFHUGB	platelet-derived g
22	1060	20.1	1048	T30815	platelet-derived g
23	998.5	19.0	1338	S09982	protein-tyrosine k
24	984.5	18.7	1336	I60598	platelet-tyrosine k
25	983.5	18.7	1333	I78875	receptor tyrosine
26	980	18.6	1379	JC4954	vascular endotheli
27	970	18.4	1330	A5010	gag-kit polypeptid
28	967.5	18.4	790	F0MVH2	vascular endotheli
29	965	18.3	1348	S51656	

ALIGNMENTS

RESULT 1

A39931
protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 04-Feb-2000
C:Accession: A39931
R:Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.
Cell 65, 1143-1152, 1991
A:Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell
A:Reference number: A39931; MUID:91292518; PMID:1648448
A:Accession: A39931
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-992 <NAT>
A:Cross-references: GB:M64689; NID:g193327; PIDN:AAA37634.1; PID:g193328
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein
F:609-953/Domain: protein kinase homology <KIN>
F:617-625/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 5264; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 4.7e-262;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

30	954.5	18.1	1298	2	A48999	protein-tyrosine k
31	954.5	18.1	1356	2	JC1402	protein-tyrosine k
32	952	18.1	1363	2	I58375	protein-tyrosine k
33	926.5	17.6	1367	2	A41228	protein-tyrosine k
34	822	15.6	160	2	A39061	protein-tyrosine k
35	815	15.5	813	1	A49123	fibroblast growth
36	810.5	15.4	823	2	B35963	protein-tyrosine k
37	798	15.2	797	2	S38579	fibroblast growth
38	791	15.0	821	1	TVHUF2	fibroblast growth
39	791	15.0	824	2	S24108	protein-tyrosine k
40	789	15.0	806	2	A35963	protein-tyrosine k
41	788.5	15.0	769	2	S16236	fibroblast growth
42	788.5	15.0	822	2	A45081	fibroblast growth
43	780.5	14.8	822	2	A41794	keratinocyte growt
44	779.5	14.8	457	2	S44269	platelet-derived g
45	777	14.8	821	1	TVMSBK	fibroblast growth


```
QY 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLSWTWKCKSDKSPNCTEIEIP 480
|||||
DB 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLSWTWKCKSDKSPNCTEIEIP 480
QY 481 EGVNKKANKRVFGQWSSSTLNNSBAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
|||||
DB 481 EGVNKKANKRVFGQWSSSTLNNSBAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
QY 541 DNISFYATIGLCPLFFIVLVILVICHYKQFRYESOLOMIOVTPGLDNEYFYVDFRDYIEY 600
|||||
DB 541 DNISFYATIGLCPLFFIVLVILVICHYKQFRYESOLOMIOVTPGLDNEYFYVDFRDYIEY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
|||||
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFKHRTWTEIFK 720
|||||
DB 661 SELKMMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFKHRTWTEIFK 720
QY 721 EHNFSYPTFOAHNSNSMPGSRREVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
|||||
DB 721 EHNFSYPTFOAHNSNSMPGSRREVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
QY 781 EDNLVLTFFEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
|||||
DB 781 EDNLVLTFFEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSYVVRGNARLPVKWMAPESELEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
|||||
DB 841 SDSYVVRGNARLPVKWMAPESELEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEQPFYATGEGYIFVMSQWAFDSKRPSPQVKIHRERS 992
|||||
DB 901 FYKLIOGFKMEQPFYATGEGYIFVMSQWAFDSKRPSPQVKIHRERS 992
QY 961 SIHLPKQAAPQORGLRAQSPQVKIHRERS 992
|||||
DB 961 SIHLPKQAAPQORGLRAQSPQVKIHRERS 992

RESULT 2
Flt3 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C:Accession: S18827
R:Rosnet, O.; Marchetto, S.; deLapeyriere, O.; Birnbaum, D.
A:Title: Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1
A:Reference number: S18827; MUID:92019834; PMID:1656368
A:Accession: S18827
A:Molecule type: mRNA
A:Residues: 1-1000 <ROS>
A:Cross-references: EMBL:X59398; NID:g50978; PID:CAA42041.1; PID:g50979
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C:Keywords: ATP
F:609-953/Domain: protein kinase homology <KIN>
F:617-625/Region: protein kinase ATP-binding motif

Query Match 96.9%; Score 5102; DB 2; Length 1000;
Best Local Similarity 96.7%; Pred. No. 9.2e-254;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
QY 1 MRALAQSDRRLLLVLSVMILETVTNQDLPVTKCVLISHENNSSAGKPSYRMVRGS 60
|||||
DB 1 MRALAQSDRRLLLVLSVMILETVTNQDLPVTKCVLISHENNSSAGKPSYRMVRGS 60
QY 61 PEDLOCTPRQSEGTVEATVEAEGSITLQVQLATPGDLSCLVWFVKHSSLCQPHFD 120
|||||
DB 61 PEDLOCTPRQSEGTVEATVEAEGSITLQVQLATPGDLSCLVWFVKHSSLCQPHFD 120
QY 121 LQNRGIVSMAILNVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLVYLRPRPKMENQ 180
|||||
```

```
DB 121 LQNRGIVSMAILNVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLVYLRPRPKMENQ 180
|||||
QY 181 DALLCISEGVPEPTVEMVLCSSSHRESCKEGPAVVRKEEKVLHFGTDIRCCARNALGR 240
|||||
DB 181 DALLCISEGVPEPTVEMVLCSSSHRESCKEGPAVVRKEEKVLHFGTDIRCCARNALGR 240
QY 241 ECTKLFITIDLNQAPQSTLPOLFLKVGEPWIRKAIHVNHGFGLTWELEDKALEEGSYFE 300
|||||
DB 241 ECTKLFITIDLNQAPQSTLPOLFLKVGEPWIRKAIHVNHGFGLTWELEDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVITILEKGFINATSSQBEY 360
|||||
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVITILEKGFINATSSQBEY 360
QY 361 EIDPYEFCFSVRKAYPRIRCTWIFSOAFPPCQORGLDGYISKFCDHKNKPGYEIFY 420
|||||
DB 361 EIDPYEFCFSVRKAYPRIRCTWIFSOAFPPCQORGLDGYISKFCDHKNKPGYEIFY 420
QY 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLSWTWKCKSDKSPNCTEIEIP 480
|||||
DB 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLSWTWKCKSDKSPNCTEIEIP 480
QY 481 EGVNKKANKRVFGQWSSSTLNNSBAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
|||||
DB 481 EGVNKKANKRVFGQWSSSTLNNSBAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
QY 541 DNISFYATIGLCPLFFIVLVILVICHYKQFRYESOLOMIOVTPGLDNEYFYVDFRDYIEY 600
|||||
DB 541 DNISFYATIGLCPLFFIVLVILVICHYKQFRYESOLOMIOVTPGLDNEYFYVDFRDYIEY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
|||||
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFKHRTWTEIFK 720
|||||
DB 661 SELKMMTHLGHHDNIYNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFKHRTWTEIFK 720
QY 721 EHNFSYPTFOAHNSNSMPGSRREVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
|||||
DB 721 EHNFSYPTFOAHNSNSMPGSRREVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
QY 781 EDNLVLTFFEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
|||||
DB 781 EDNLVLTFFEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSYVVRGNARLPVKWMAPESELEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
|||||
DB 841 SDSYVVRGNARLPVKWMAPESELEFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEQPFYATGEGYIFVMSQWAFDSKRPSPQVKIHRERS 992
|||||
DB 901 FYKLIOGFKMEQPFYATGEGYIFVMSQWAFDSKRPSPQVKIHRERS 992
QY 961 SIHLPKQAAPQORGLRAQSPQVKIHRERS 992
|||||
DB 961 MGVNPEHPSIYQNRRLPSREAGS-EPPSPQAQVKIHRERS 1000
|||||

RESULT 3
A36873
protein-tyrosine kinase (EC 2.7.1.112) STK-1 precursor - human
N:Alternate names: stem cell tyrosine kinase 1
C:Species: Homo sapiens (man)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 04-Feb-2000
C:Accession: A36873
R:Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; B
Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994
A:Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+)
A:Reference number: A36873; MUID:94119906; PMID:7507245
A:Accession: A36873
A>Status: preliminary
```

A:Molecule type: mRNA
A:Residues: 1-993 <SMA>
A:Cross-references: GB:U02687
A:Note: in the authors translation, an additional residue Ala is shown after 420-Ala and
C:Genetics:
A:Map position: 13q12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-speci
F:608-950/Domain: protein kinase homology <KIN>
F:616-624/Region: protein kinase ATP-binding motif

Query Match 83.7%; Score 4408.5; DB 2; Length 993;
Best Local Similarity 84.0%; Pred. No. 2.8e-218;
Matches 836; Conservative 57; Mismatches 89; Indels 13; Gaps 6;
QY 1 MRALAQSDRRLLLVLSVILETVTQDLPVVKVLIHSHENNGSSAGKSPSSYRWGRGS 60
DB 1 MPALA-RDAGTVPLLVFSAFIIGTITNODLPVVKVLIHSHENNGSSAGKSPSSYRWGRGS 59
QY 61 PEDLQCTPRQSEGTVEAATVEVAESGITLQVOLATPGDLSCLWVFKHSLGCGPHFD 120
DB 60 PEDLGCALRPOSGTVEAATVEVAESGITLQVOLATPGDLSCLWVFKHSLGCGPHFD 119
QY 121 LQNRGIYSMAILNVTQAGEYLHIIQSERANYTLFTVNVVRDQOLYVLRPRYFKMENQ 180
DB 120 LQNRGVVSWILKMTQAGEYLHIIQSEATNYTLFTVNVVRDQOLYVLRPRYFKMENQ 179
QY 181 DALICISEGPEPTVENVVLCSSHRESCKEKPAPVVRKEEVHLHFGDTRCCARNALGR 240
DB 180 DALVCISEGPEPTVENVVLCSSHRESCKEKPAPVVRKEEVHLHFGDTRCCARNALGR 239
QY 241 ECTKLFTDLQAQOSTLPOLFLKVGPELWIRCKRAIHNHGHGFLTWELKALEGSGYFE 300
DB 240 ECTRLFTIDLQNOTFTLLPQLFLKVGPELWIRCKRAIHNHGHGFLTWELKALEGSGYFE 299
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTSSSKHPSQSALVTILEKGFINATSSQBEY 360
DB 300 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTSSSKHPSQSALVTILEKGFINATSSQBEY 359
QY 361 EIDPYEFCFSVRKAPRIRCTWIFQASPCBORGLEDGYSTKFCDHKNKGEYIFY 420
DB 360 EIDQYEEFCFSVRKAPRIRCTWIFQASPCBORGLEDGYSTKFCDHKNKGEYIFY 419
QY 421 -AENDDAQFTKMTLNIRKRPQVLANASASQSSDGYPLPSWTWKKCDKSPNCTEEI 479
DB 420 -AENDDAQFTKMTLNIRKRPQVLANASASQSSDGYPLPSWTWKKCDKSPNCTEEI 478
QY 480 PEGVNRKANRKFQGVVSSSTLNKSEAGKGLLVKCCAYNSMGTSCEIFLNSPGPPFFI 539
DB 479 TEGVNRKANRKFQGVVSSSTLNKSEAGKGLLVKCCAYNSMGTSCEIFLNSPGPPFFI 538
QY 540 QDNISFYATIGLCLPFIWLVILVILCHYKKQFRYESQIQMIVTGLDNEYFYVDFRDE 599
DB 539 QDNISFYATIGLCLPFIWLVILVILCHYKKQFRYESQIQMIVTGLDNEYFYVDFRDE 598
QY 600 YDLKWEPPRENLEPGKVLGSCAFGRVNNATAYGSKTGVSQVAVKMLKADSCKEAL 659
DB 599 YDLKWEPPRENLEPGKVLGSCAFGRVNNATAYGSKTGVSQVAVKMLKADSCKEAL 658
QY 660 MSELKMTLGHHDNIIVNLGACFTLSGPVYLIFFECVCGDILLNLRKREKFRHTWTIF 719
DB 659 MSELKMTLGHHDNIIVNLGACFTLSGPVYLIFFECVCGDILLNLRKREKFRHTWTIF 718
QY 720 KEHNFSSYPTFOAHSSNMPGSRVQLHPPLDQLSGFGNGSIHSEDETEYENQKRLABEE 779
DB 719 KEHNFSSYPTFOAHSSNMPGSRVQLHPPLDQLSGFGNGSIHSEDETEYENQKRLABEE 776
QY 780 EEDLNVLTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTHTGKVKICDFGLARDI 839
DB 777 EEDLNVLTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTHTGKVKICDFGLARDI 836
QY 840 LSDSSVYVRGNARLPVKWMAPELFGIYTIKSDVWSYIGILLWEIFSLGVNYPGIPVDA 899
DB 837 MSDSNVYVRGNARLPVKWMAPELFGIYTIKSDVWSYIGILLWEIFSLGVNYPGIPVDA 896
QY 900 NYFKLIQSGFKWEDPPYATGEGYFVWQSWAFDSRRRPSFPNLTSLGCOLAEAEAC-- 957
DB 897 NYFKLIQSGFKWEDPPYATGEGYFVWQSWAFDSRRRPSFPNLTSLGCOLAEAEAMQ 956
QY 958 -----IRTSIHLPKQAAPQORG-GLRAQSPQORVK 986
DB 957 NVDRGVSECPHTYQNRPRPSREMDLGLLSPQAQVE 991

RESULT 4
TVCTWD
macrophage colony-stimulating factor 1 receptor precursor - cat
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 13-Jun-1997
C:Accession: A31636
R:Woolford, J.; McAlliffe, A.; Rohrschneider, L.R.
Cell 55, 965-977, 1988
A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are requ
A:Reference number: A31636; MUID:89077553; PMID:2849512
A:Accession: A31636
A:Molecule type: mRNA
A:Residues: 1-980 <WOO>
A:Cross-references: EMBL:X03663
C:Genetics:
A:Gene: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming prote
fic protein kinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <EXT>
F:24-509/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-240/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-484/Domain: immunoglobulin homology <IMM5>
F:510-534/Domain: transmembrane #status predicted <TM>
F:535-980/Domain: intracellular #status predicted <INT>
F:577-915/Domain: protein kinase homology <KIN>
F:585-593/Region: protein kinase ATP-binding motif
F:42-84,124-177,224-278,417-482/Disulfide bonds: #status predicted
F:45,73,94,153,275,303,335,410,477,490/Binding site: carboxylate (Asn) (covalent) #s
F:613,630,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.9%; Score 1258.5; DB 1; Length 980;
Best Local Similarity 32.7%; Pred. No. 4.3e-57;
Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;
QY 75 TVYEAAATVEAESGISTLQVOLATPGDLSCL-----W---VFKHSLGCGPHFDLQNRG 125
DB 13 TAWHAQGVVQPSGPELVVEPGTTVTLRCVGVSVEMDGPISPHNLDLDPSSI---- 68
QY 126 IVSMAILNVTQAGEYLHII-----OSERANYTLFTVNVVRDQOLYV---RRPY----- 173
DB 69 ---LTNNATONTCTY---HCTEPGNPGGNATI-----HLVKKDAPRWKVLQAE 114
QY 174 FRKMENQDALL-CISEGVPEPTVENVVLCSSHRESCKEKPAPVVRKEEVHLHFGDTR-- 230
DB 115 VTVLEGQDALLPCL---LTDPALE-----AGVSLVRVRGRPVLRQTNYSFSPWHGFTIHK 166
QY 231 -----RCCARNALGRECTKL-----FTDILNQAPOSTL-PQLFLKV-GEPLWIRCK 274
DB 167 AKFIENHYQCSAR-VDRGTVTSMGIWLKVQKDISGPATLTLEPAELVRIQGEAAQIVCS 225
QY 275 AIHVNHGFLTWELKEDKALEGSGYFEMSTYSNRTMIRILLAFVSSVGRNDTGYYTCSSS 334
DB 226 ASNDVNFEDVSLRHGDTKL---TISQSDFDHNRQY-KVLTNLNDHVSQDAGNYSCTAT 281
QY 335 K---HPSQALVTLIEKGFINATSSQE-EYEDPYEFCFSVRKAPRIRCTW-I----- 385

Db 282 NAMGNHSAAMFVRFVESAYLNLTSEOSLLQEVTVGKVDLOVKVEAYPGLESFNWYLG 341
Qy FSOASFPCEQRGLEGDYSISKFCD---HNKNPGEIFYAENDDAQFTKMTFLNIRKKPQ 441
Db 342 FSDYDKLDFVTIKDYRTSTLSPLRKRSEAGRSFLARNAGGQNALTFFELTLPPE 401
Qy 442 VLANASAQAS-----CSSDGYPPLPSWTWKCKSKDNCTEE---IPEGVMNKANKRVFG 494
Db 402 VRVMTLINGSPTLLCEASGYPOPSVTWVOCRSHTRDCDESAGLVLEDSSHSEVLVSQVPFH 461
Qy 495 QWSSSTINMSAGKGLLVKVCAYNSMGTSCTEILNSPGPFPIODNISFYATIGLC-- 552
Db 462 EVIVHSLLAIGTLNHRTTECFRAFNSVGNSQTWPISIGAHTQLDPBELTFPVLLTWS 521
Qy 553 -LPETVLIVLTCHRYKKOFRYESQIQVTVTGPDONEYFYVDRDYEDLKWEFFPRENL 611
Db 522 IMALLLLLLLLLXYKOKPKYVRWKIIE-SYEGNSYTFIDPTQLPYNEKEWFPRNNL 579
Qy 612 EFGKVLGSAGAFQVWNATAYGISKGTGVSTQAVKMUKAKADSCKEALMSELKMMTHLGH 671
Db 580 QFGKTLAGAGFKGVYEATAFGLGKEDAVLKVAVKMLKSTAHADEKALMSELKIMSHLGQ 639
Qy 672 HDNVILLGACTLSGPVYLIFCYCCYCDLLNYLRKRE-----KF 711
Db 640 HENIVNLLGACTHGFPVLVITECYCGDLLNFLURQAAMLGPSLSVGQDPGAGAGYKNI 699
Qy 712 HRTWEIFEKHFNFSPYTFQAHNSSMPGSRREVOLHPPLDQLSGFGNSIHSEDEIEYEN 771
Db 700 HLEKYVRDSDFS-----SOGVDYTYVMRPVSTSSNDSEEDL---- 740
Qy 772 QKRLAEEDLNVITFDLLCFAYOVAKGMFELEPKSCVHRDLAARNVLVTHGKVVKIC 831
Db 741 -----CKEDGRPLELRDLLHFSQVAQMAFLASKNCIHRDVAAARNVLLTSGRVAKIG 793
Qy 832 DFGLARDILSDSYVVGRNARLPVKWMAPELSEGIYTIKSOWSYSIGILLWEIFSLGVNP 891
Db 794 DFGLARDIMNDSNYIVKGNARLPVKWMAPESIDCVYTIVOSDWWSYGILLWEIFSLGLNP 853
Qy 892 YPGIPVDANFYKLIOGGFKMEOPFYATEGIYFMVQSCWAFDSRRKRPFPNLTSLFGCOLLA 951
Db 854 YPGILVNSEFYKLVRDGYOMAQPAFAKPNIIYSIMQACWALETPTRRPTFOQICSLL-QKQ 911
Qy 952 EAEEACIRTSHLPKOAAPOQ-----RGLRAQSPOKQ 984
Db 912 AQEDRRVPNTNLPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 953

RESULT 5
TMVMD
protein-tyrosine kinase (EC 2.7.1.112) fms precursor - feline sarcoma virus (strain McDe
C;Species: feline sarcoma virus
A;Note: host Felis sp. (cat)
C;Date: 27-Nov-1985 #sequence_revision 31-Dec-1991 #text_change 13-Jun-1997
C;Accession: A00654
R;Hampe, A.; Gobet, M.; Sherr, C.J.; Galibert, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 85-89, 1984
A;Title: Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected ho
C;Reference number: A00654; MUID:84119469; PMID:6582485
A;Accession: A00654
A;Molecule type: DNA
C;Residues: 1-941 <HAM>
C;Comment: This protein is synthesized as a gag-fms polyprotein.
C;Genetics:
A;Gene: fms
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
otin kinase
F;1-23/Domain: signal sequence.#status predicted <SIG>
F;24-941/Product: protein-tyrosine kinase fms #status predicted <MAT>
F;24-509/Domain: extracellular #status predicted <EXT>
F;35-86/Domain: immunoglobulin homology <IMM1>
F;120-179/Domain: immunoglobulin homology <IMM2>

F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-484/Domain: immunoglobulin homology <IMM5>
F:510-534/Domain: transmembrane #status predicted <TM>
F:535-941/Domain: intracellular #status predicted <INT>
F:577-915/Domain: protein kinase homology <KIN>
F:585-593/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted
F:45,73,94,153,275,286,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent)
F:613,630,776/Active site: Lys, Glu, Asp #status predicted
F:781/794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 33.7%; Score 1247; DB.1; Length 941;
Best Local Similarity 33.2%; Pred. No. 1.6e-56;
Matches 325; Conservative 162; Mismatches 347; Indels 144; Gaps 31;

Qy 75 TVYEAAATVEVAESGSIITLOVATPGDLSCL-----W---VFKHSSILGCPHFEDLQNRG 125
Db 13 TAWHAQGVVLPQSGPELVPEPGTVILRCVNGSVENDGPISPHNLDLDPSSI----- 68
Qy 126 IVSMAILNVETQAGEYLLHI-----QSERANYTVLFTVNVRTQLYVL--RRPY----- 173
Db 69 ---LTTNNATFQNTGY--HCTEPCNPRGGNATI-----HLXKDPARPKNKVLQAE 114
Qy 174 FRKMNQDALL-CTSGVPEPTVFWLCCSHRESCKEKGPAVVRKEKVLHLEFGTDI-- 230
Db 115 VTVLEGODALLPCL---LTDPALE-----AGVSLVRVRGPRVLRQTNYSFSPWGHGFTIHK 166
Qy 231 -----RCCARNALGRECTKL-----FTIDLNOAQPSLT-POLFVKV-GEPLWTRCK 274
Db 167 AKFTENHYVQCSAR-VDGRTVTSGIWLKVKOKDISGPATLTLEPAELVRIQGEAAIVCS 225
Qy 275 AIHNHGFGFLWELEDKALEEGSYFEMSTYSTRNTRMIRILLAFYSSVGRNDTGYYTCSSS 334
Db 226 ASNIDVNFDSLVRHGDTKL---LTSQOSDFHNRQY-KVLTLNLDHVSFQDAGNYSCTAT 281
Qy 335 K---HPSQSALVTILEKGFINATSOE-EYELDPYKEFCFSVRKAPRIRIC-TWI----- 385
Db 282 NAWGNHSA5MFRVYESAYSNLTSQESLLQEVTVGKEVDLQVKVEAYPGLESFNNTYLG 341
Qy 386 FSOASFCEQRLGEDGYSISKFCD-----HKNKPEYIFYAENDDAQFTKMTFLNIRKKPQ 441
Db 342 FSDYQDKLDFVTIKDTRYTTLSLPLRKRESGRYSFLARNAGGONALTTELRLRYPPE 401
Qy 442 VLANASASQAS-----CSSDGYPLPSWTWKKCKSDKSPNCTEE---IPEGWNKANKRKVEG 494
Db 402 VRVMTLLINGSDTLLCEASGYQPQSVTWQCRSHTRDCEAGLVLEDSSHSEVL5QVPFY 461
Qy 495 QWVSSTLNMSEAKGLLVKCCANSMGTSETFLNSPGPFPTQDINISFYATIGLC-- 552
Db 462 EYVHSLLAIGTLEHNRNTRYECRAFNSVGNSSQTFWPISIGAHTPLPDELFTPVLLTCS 521
Qy 553 -LPETVVLVLILCHYKKOFYESOLOMIOVTGPLDNEYFYVDFRDYEDLKWEPREN 611
Db 522 IMALLULLLLLLLYKKPKYQVRWKIIE--STEGNSYTFIDTQLPYNKEKWEPRNLL 579
Qy 612 EFGKVLGSGAFGRVYNATAYGISKTGVSIOAVKMLKEKADSCKEALMSLKMHTHLL 671
Db 580 QFGKTLGTGAFGKVVVEATAFGLGEDAVLKVAVKMLKSTAHADEKALMSLKMHLGQ 639
Qy 672 HDNIVNLLGACTLGSPPVLIFFEYCYGDLNLYLRSKRE-----KF 711
Db 640 HENIVNLLGACTHGGPVLVITFEYCCYGDLNLFRRQAEAMPGPSLSVGQDPEAGAGYKNI 699
Qy 712 HRTWTEIFEKHNFSYPTFQAHSNMPGCSREVQLHPDLQDSGFNGNSIHSEDEIEYEN 771
Db 700 HLEKKYVRD5GS-----SGQVDTYVEMRPVSTSSSND5FSEEDL----- 740
Qy 772 OKRLAEEEDLNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLYTHGKVVKIC 831
Db 741 -----GREGRPELRLDLHFSSQVAAQWAF LASKNCIHRDVAARNVLYTHSGRVAIG 793
Qy 832 DFGLARDILSDSSYYVVRGNARLPVKWMAPE5LIFEGIYTIKSDVMSYGTLLWEIFSLGVNP 891

A:Accession: 159083
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 874-972 <R3>
 A:Cross-references: GB:M14193; NID:g182521; PIDN:AAA35834.1; PID:g182522
 R:Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.
 Cell 42, 421-428, 1985
 A:Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del
 A:Reference number: 152772; MUID:85282599; PMID:4028159
 A:Accession: 152772
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 244-295 <R4>
 A:Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423
 C:Genetics:
 A:Gene: GDB:CSFIR; FMS
 A:Cross-references: GDB:120600; QMIM:164770
 A:Map position: 5q33.2-5q33.3
 A:Introns: 1771; 10371; 19871; 24373; 29771; 36172; 40071; 44072; 50471; 54273; 58571; 6
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 fic protein kinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAT
 F:24-512/Domain: extracellular #status predicted <EXT>
 F:35-86/Domain: immunoglobulin homology <IMM1>
 F:120-179/Domain: immunoglobulin homology <IMM2>
 F:217-280/Domain: immunoglobulin homology <IMM3>
 F:316-383/Domain: immunoglobulin homology <IMM4>
 F:412-487/Domain: immunoglobulin homology <IMM5>
 F:513-537/Domain: transmembrane #status predicted <TMM>
 F:538-972/Domain: intracellular #status predicted <INT>
 F:580-917/Domain: protein kinase homology <KIN>
 F:588-596/Region: protein kinase ATP-binding motif
 F:42-84, 127-177, 224-278, 419-485/Disulfide bonds: #status predicted
 F:45, 73, 153, 240, 275, 302, 335, 412, 428, 480/Binding site: carbohydrate (Asn) (covalent)
 F:616, 633, 778/Active site: Lys, Glu, Asp #status predicted
 F:783, 796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1221.5; DB 1; Length 972;
 Best Local Similarity 32.5%; Pred. No. 3.3e-55;
 Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;
 QY 76 VYEATVEAESGSIITLOVOLATPGDLSCLVWFKHSSLCGCPHFLQNRGIVSMAIL---132
 Db 22 VIESVPELVKPGAT--VTLRCVNGSVENDGPPS-----PHWTLYSDG--SSILSTN 72
 QY 133 NVTTQAGEYLLHIOSEIRANTVLTVMVNDTQTLVLRPPY-----FRKMNQDALL-C 185
 Db 73 NATFQNTGTYRCTEPGDPGLGSAATHLYVKDP-----ARPWNVAQVVFEDQALLPC 127
 QY 186 ISEGVPEPTVEWVLCSSHRECKEGPAPVVKEEK-----VLHE---LF 226
 Db 128 L---LTPVL-----EAGVSLVRVRGRLMRHTNYSFSPWHGETIHRAKFIQ 171
 QY 227 GTDIRCCARNALGRECTKLTIDLNOAQSTLP-----QLFLKVGPELWIRCKAI 276
 Db 172 SODYQCSALMG-GR---KWSISLTKVQKVIIPGPPALTLPALVRLRIRGEAAIVCSAS 227
 QY 277 HVNHGFGLTWELEKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSS-- 334
 Db 228 SDVNVDFVLQHNNTKL--AIPQSDPHNNRYQ-KVLTNLDQVDFQAGNYSVASNV 283
 QY 335 --KIPSQSAVLTILEKGINATSSQE-EYEDIDPEKFCFVSFRKAYPRIR-CTWIFSOAS 390
 Db 284 QGKH-STSMFFRVVSEAYLNLSEQNLIQEVTVGEGNLKVMVEAYPGLOGFNWY---- 338
 QY 391 FPCBQRLGLEDGYSISKFCDHKNKP-----GEYIFYAEN 423
 Db 339 -----LGPFSHQPEKPLANATTKDVTYRHTTILSLPKLPSEAGRVSLARN 385
 QY 424 DDAOFTKMTLNIKKPOV-----LANASASQASCSGYPPLPSMTWKKCKSDKSPNCTEE 478

RESULT 9
 I51703
 C-kit-related kinase 1 (Xkrk1) - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
 C:Accession: I51703
 R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.
 Mech. Dev. 50, 217-228, 1995
 A:Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem
 A:Reference number: I51703; MUID:95344996; PMID:7619732
 A:Accession: I51703
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-954 <BAK>
 A:Cross-references: EMBL:248770; NID:g763033; PIDN:CAA88688.1; PID:g763034
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
 C:Keywords: ATP
 F:575-915/Domain: protein kinase homology <KIN>
 F:583-591/Region: protein kinase ATP-binding motif
 Query Match 23.2%; Score 1220; DB 2; Length 954;
 Best Local Similarity 32.0%; Pred. No. 3.9e-55;
 Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;
 QY 77 YEATVEAESGSIITLOVOLATPGDLSCLVWFKHSSLCGCPHFLQNRGIVSMA 130
 Db 15 YTGDAVPKINDGDRVTVMVNDKVSLECDRAHLVTAFQKSLMKKPR-DLKSRLP----- 68
 QY 131 ILNVTET-----QAGEYLLHIOSEIRANTVLTVMVNDTQTLVLRPPYFKME 178
 Db 69 -LNNSETDOFFVIKADLRHIGRYICTNTQENTSV--SLFKVDPAPELDPIDFIDVTE 125
 QY 179 NODAL-LCISEGVPEPTVEWVLCSSHRECKEGPAPVVKEEKVHLFETDTRC----- 232

Db 126 GADTVGMCFF---TDPDMBIAI-----EKC---DGSPLPE-----NFTTTDIEAGITIK 169
QY 233 -----CARNALGR-ECTKLTIDLNQAPQSTLPLQLF-----KYCEPLWIRC 273
Db 170 TVOLAFDSCYVCGNKGSGVTKSSSTFHSVHKVPKK-VPTVFLSKSRQLVKGEPEVTC 228
QY 274 KAIHNVHGFGLTW-ELEDKALEEGSVFENSTYNTNTRMIRILLAFVSSVGRNDTGYTCS 332
Db 229 AVLDFSVTKAQWLDVKEGVTQANFRSSNVFSLTKS-----DGVPSERSTFTTCQ 282
QY 333 SKKHPQS---SALVTILEKGFNATSSQE-EYEIDPYEKFCFSVRPKAYPR-IRCTWIFS 387
Db 283 AENAIQOVNATFTLDVIDYGVNLTVALENTTISVAGNDLNLVKVVIDAYPHPDGQVWYTF 342
QY 388 QASFCEQORGLDGYISIKFCDHK-----KGEYIFYAENDDAQTKMF 432
Db 343 NETL-----LNTSDHYVATK-DEGNRRVYSELHLRLKTEKGVYTFYTNSSDDASVSF 396
QY 433 TLNIRKKQVLANASASQAS---CSSDGYPLPSWTWKKCDKSPNCTEETPEEGVWKK--A 488
Db 397 NIQVTRPEILLARTSEGTLOCVATGFPVPAIQWYFCPSQORCTDYPPLSPVNEKFTQ 456
QY 489 NRKVFQWSSSTLNMSEAGKLLVKCCAYNSMGTS-----CETIFLNSPGP 535
Db 457 ENSSLGRIVVESTIDVNDLKNKGTVCQVASNEVESAYSFSAIKEKRLRTHLFT----- 511
QY 536 PFIQDNISFYATIGLCLPFIIVLVILVILCHKKYKQFYESQLOMI-QVTGPLDNEYFYVD 594
Db 512 -PLL---IGFIAAAGL-----MCIAVAVLMYKYLQPKYEIQMKVVEEING---NNVYID 560
QY 595 PRDYEYDLKWEPPRENLEPGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLKEKADSC 654
Db 561 PTQLPYDNKWEPPRDLRCGKILGAGAFKQVVEATAYGLLKEDSRILTAVKMLKPSAHT 620
QY 655 EKEALMSELKMTLGHHDNIIVNLGACTLSGPVYLIFCYCCYGDLLNLYLRSKREKFRHT 714
Db 621 EREALMSELKLSYLGHKHNIVNLGACTVGGPTLVITEYCCYGDLLNLYLRSKRDSEF--- 677
QY 715 WTEIFEKHNFSYPTFOAHNSMSPGSRVQLHPPLDQLSGFNGNSIHSEDELEYENOK- 773
Db 678 -----ICPKFEDNS-----EALYKNI-----LNRDMCGEMSEYIDMKP 713
QY 774 -----RLAEEDDLNVLTFEDLLCFAYQVAKGMEFLFKSC 810
Db 714 AVSYVVPKTKRRKSGSPGQDQSVSIPEDDL-ALDTEDLNFSQVAGMFLASKNC 772
QY 811 VHRDLAARNLVTHGKVVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYTI 870
Db 773 IHRDLAARNILLTHGRITKICDFGLARDIRNDSNYVVKGNARLPVKWMAPESLFHCYTF 832
QY 871 KSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATEGIFYVMQSCWA 930
Db 833 ESDVWSYGILLWEIFSLGSSYPRIQVDSKFKYKMIKIDGVRNMSPECAPLEMYEIMRSCWN 892
QY 931 FDSRKRPSFNLTSFLGCQLAAE 954
Db 893 SDPLKRPFTKQIVQWVEQQLSDSK 916

RESULT 10
I45877

protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs

C:Species: Bos primigenius (aurochs)

C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 20-Apr-2000

C:Accession: I45877

R:Kubota, T.; Hirono, H.; Sasaki, E.; Sakurai, M.

Gene 141, 305-306, 1994

A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.

A:Reference number: I45877; MUID:94215924; PMID:7512939

A:Accession: I45877

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-977 <KUB>

A:Cross-References: GB:D16680; NID:G516659; PIDN:BAA04084.1; PID:G516660
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:329-395/Domain: immunoglobulin homology <IM>
F:588-932/Domain: protein kinase homology <KIN>

Query Match 23.1%; Score 1215; DB 2; Length 977;

Best Local Similarity 32.8%; Pred. No. 7.2e-55;

Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;

QY 134 VTE----TQAGEYLLHIQSERANYTLFTVNVDRDQ-LYVLRPRPYFRKMNQDALLCI-- 186

Db 83 ITEKAEATNTGNYTC---TNKGLSSIIYVVRDPEKFLDIDLPLYKEENDTLVRCLPT 139

QY 187 -----SGVPEP-TVEWVLSSSHRESCKECPAVVRKEEKLHLELFGTDIRC-- 232

Db 140 DPEVTNYSLTGCEGKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLH 185

QY 233 CARNALGRE-CTKLTIDLNQAPQS-----TLPLQLKVGGEPLWIRCKAIHNVHGFGLT 285

Db 186 CSANQRGKSMLSKKFTLKVRRAIKAVPVSVSKTSYLLREGEFAVTCILIKDVSSVDSM 245

QY 286 WELEDKALBEGSYFEMSTYNTNTRMIRILLAFVSSVGRNDTGYTCSKSHPSQSALVTI 345

Db 246 WIKENSQOTKQATKKNSHQGDFSYLRQERLTISSARVNDGVFMVYANNTFGSANVTIT 305

QY 346 LE---KGFNA-TSSQEEYEIDPYEKFCFSVRPKAYPR-IRCTWIFSQAAPPCQORGLD 400

Db 306 LEVVDKGFINIFPMNTTVPVNDGENVDLVVEYAYPKVHRQWIYMNRT----- 355

QY 401 GYSISKFCDHK-----NKPGEYIFYAENDDAQTKMFTLNIRKK 439

Db 356 ---STDKWDYDPKSENEISIRYVNLHLRLKGTGGTYTFHVSNSDVNSSTFNVVNTK 413

QY 440 PQVLANASA---SOASCSGDYPLPSWTWKKCDKSPNCTEETI-PEGVWKKANRKYVQW 496

Db 414 PEILTHDLRVNGLQCVAAQFPEPTIDWYFCPGTEQCSVPVGPVDVQIQNSVSPGKL 473

QY 497 VSSSTLNMSEAGKLLVKCCAYNSMGTS-----CETIFLNSPGPPF 539

Db 474 VVYSTIDDSFTFKINGVTECRAYNDVGKSSASFNAFKGNSKEQIHAHTLFT-----PLL 527

QY 540 QDNISFYATIGLCPFIIVLVILVILCHKKYKQFYESQLOMI-QVTGPLDNEYFYVDRDY 598

Db 528 ---IGFVIAAGLCIFVMI----TYKYLQKPMYEVQWKVVEEING---NNVVIDPTQL 577

QY 599 EYDLKWEFPRENLEPGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLKEKADSCKEA 658

Db 578 PYDHKWEFPNRISFGKTLGAGAFKQVVEATAYGLIKSDAAMTAVAKMLKPSAHLTEREA 637

QY 659 LMSLKMTHLGHHDNIIVNLGACTLSGPVYLIFCYCCYGDLLNLYLRSKREKFRHTWTEI 718

Db 638 LMSLKVSLVSLGHNMIVNLGACTIGGPTLVITEYCCYGDLLNLYLRSKRDSEF----- 690

QY 719 FKEHNFSYPTFOAHNSMSPGSRVQLHPPLDQLSGFNGNSIHSEDEI-----EYENQ 772

Db 691 -----ICSKQEDHAEVALYK-----NLLHSKSSCNDSTNEYMDM 725

QY 773 K-----RLAEEDDLN-----VLTFEDLLCFAYQVAKGMEFL 806

Db 726 KPGVSYVVPKTKRRKSARISYTERDVTPTAIMEDDELALDLEDLSFSYQVAKGMAFLA 785

QY 807 FKSCVHRDLAARNVLVTHGKVVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEG 866

Db 786 SKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNYVVKGNARLPVKWMAPESLFNC 845

QY 867 IYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATEGIFYVMQ 926

Db 846 VYTFESDWSYGILFELWELFSLGSSPYGMPVDSKFKYKMIKIDGVRNMSPEHAPAEYDLMK 905

QY 927 SCWAFDSKRKRPSPNLTSLFLGCQLAAE 953

Db 906 TCWDADPLKRPFTKQIVOLIEKQISES 932


```
RESULT 11
S16385
macrophage colony-stimulating factor 1 receptor precursor - rat
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSF-1R
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Jun-2000
C:Accession: I60321; S16385
R:Borycki, A.G.; Guillier, M.; Leibovitch, M.P.; Leibovitch, S.A.
Growth Factors 6, 209-218, 1992
A>Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis and
A:Reference number: I60321; MUID:93001225; PMID:1389227
A:Accession: I60321
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-978 <RES>
A:Cross-references: EMBL:X61479; NID:957543; PIDN:CAA43706.1; PID:957544
A>Note: in Genbank entry RRGSL1, release 113.0, the source is designated as Rattus rattu
A>Note: submitted to the EMBL Data Library, August 1991
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-relat
protein; tyrosine-specific protein kinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAT
F:20-515/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-485/Domain: immunoglobulin homology <IMM5>
F:516-535/Domain: transmembrane #status predicted <TM>
F:536-978/Domain: intracellular #status predicted <INT>
F:578-915/Domain: protein kinase homology <KIN>
F:586-594/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status
F:614,631,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.0%; Score 1210; DB 2; Length 978;
Best Local Similarity 33.5%; Pred. No. 1.3e-54;
Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;

Qy 78 EAATVEVAESGILQVQLATPGLDLSCLWFKHSLGQCPHFDLQNRGIVSMAILNVTET 137
Db 36 ETVTLCRVNSGVEMD-----GPISYWTLDPSGSGS-----TLTTRNATFK 77
Qy 138 QAGEY-LIHQISERANYTVLFTVNVRRDQLVY-----LRRPYFRKMNODALL-CIS 187
Db 78 NTGYRCTEEDPAGSTTI-----HLVYKDPAHWNLLAQEVTVVGEQEAIVPCL- 128
Qy 188 EGVPEPTVEWVLCSSHRESCKEEGPAVVRK-----BEKVHLHFGTDIRCC 233
Db 129 --ITDPALK-----DSVSLMREGGROVLKTVYFFSANRGFTIIRAKVL-----DSNTYVC 177
Qy 234 ARNALGRECTKL-FTIDLN-----QAPQSTL-POLELVK-GEPLWTRCKAIHVNHGFLTW 286
Db 178 KTVYNGRESTGTGVLKLVNRVHPEPPQIKLEPSKILVRIRGEAAQVCSATNAEVEFNIL 237
Qy 287 ELEKALE--EGSVFEMSTYSTNRTMIRILLAFVSSVSGRNDGYTC-SSSKHPQSALV 343
Db 238 KRGTOKLEIPNSQFQDNYKKVRL-----SLNAVDFQAGIYSCVASNDVGRFTATM 291
Qy 344 T--ILEKGFINATSQE--EYEDIDYKFCFSVRFKAYPRIR-CTWIFSOAPSPPQORGL 399
Db 292 NFQVVEAYNLNTSEQLLOEVSVCDSILITVHADAYSIQHNWYTLGPFE-DQKLE 350
Qy 400 -----DGYSISKPCDH--KNKPGEYIFYAENDDAQFTKFTLNRKPKQVLANASAQ 450
Db 351 FITQRIAYRYFTKFLNRVKASEAGQYFLMAQKAGWNLLITFELTLRPPEVSVTWMPVN 410
Qy 451 AS-----CSSDGYPLPSWTWKCKSKSPNCTEIPGVVN-----KKANKRVQGVSSSTL 502
```

```
Db 411 GSDVLFCDVSGYPOPSVTWMECRGHTDRCDQAQALQVWNTDHPVLSOKPDKVILQS 470
Qy 503 NMSEAGKGLLVKCCAYNSMGTCTETIFLNSPGPPFFITQDNISFYATIGLC---LPFIVVL 559
Db 471 PIGTLKHNMTYFCKTHNSVNSGSOYFRAVSLGOSKQLPDESLETPVVVACMSVSLVLL 530
Qy 560 IVLICHYKKOFRESOLOMIO-VTGPLDNEYFVDFRDYEDLKWEFEPRENLEFGKVLG 618
Db 531 LLLLLYKOKPKIQVWKIIERYEG---NSYTFIDTQLPYNEKWEFPRNNLOFGKTLG 587
Qy 619 SGAFGRVMNATAYGISGTGVSIVGKMLKEKADSKSEKEALMSSELKMTHLGHHDNIVNL 678
Db 588 AGAFKGVVETATFGLKEDAVLKAVKMLKSTAHADEKEALMSSELKMTHLGHHDNIVNL 647
Qy 679 LGACTLSGPVYLIFEYCCYGDLLNLYRSKRE-----KPHRTWTEI 718
Db 648 LGACTHGGPVLVITEYCCYGDLLNLYRSKRE-----KPHRTWTEI 707
Qy 719 FREHNFSSYPTFOAHSNNSMPGSRVQLHPPLDOLSGFNGNSIHSEDEIEYENKRLAEE 778
Db 708 RRDSGFS-----SQGVDTYVEMRPVSTSSDSFFKQD-----L 740
Qy 779 EREDNLVLTFFEDLLCFAYOVAKMEFEKSCVHRDLAARNVLVTHGVVVKICDFGLARD 838
Db 741 DKPSRPLELDLWLLHFSQVQAQMAFLASKNCIHRDVAARNVLTSGHVAKIGDFGLARD 800
Qy 839 ILSDSYVVRGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEITFSLGVNPPYGPV 898
Db 801 INDSNVVVGKGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEITFSLGVNPPYGP 860
Qy 899 ANFYKLIQSGFMEQPFYATEGIYFVMSQWAFDSKRKPSFPNLTSEL 946
Db 861 NKFYKLVKDGQYMAQPVFAPKNIYSIMQSCWDLPEPTRRPTFQOICFLL 908
```

```
RESULT 12
TVMSMD
macrophage colony-stimulating factor 1 receptor precursor - mouse
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csf1r/fms
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Jun-2000
C:Accession: S01880
R:Rothwell, V.M.; Rohrschneider, L.R.
Oncogene Res. 1, 311-324, 1987
A>Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.
A:Reference number: S01880; MUID:88217329; PMID:2966922
A:Accession: S01880
A:Molecule type: mRNA
A:Residues: 1-976 <ROT>
A:Cross-references: EMBL:X06368
C:Genetics:
C:Gene: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming prote
fic protein kinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted <
F:20-515/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-485/Domain: immunoglobulin homology <IMM5>
F:516-535/Domain: transmembrane #status predicted <TM>
F:536-976/Domain: intracellular #status predicted <INT>
F:578-914/Domain: protein kinase homology <KIN>
F:586-594/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #stat
F:614,631,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted
```

Query Match					
Best Local Similarity 22.8%; Score 1198.5; DB 1; Length 976;					
Matches 317; Conservative 148; Mismatches 347; Indels 117; Gaps					
QY	78	EAAATVEAEGSGITLQVQLATPGDLISCLLVFVKHSSILGCPHFEDLNQRGIVSMALNVTET	137	:	:
Ddb	36	EVTVLRCSVNGSVEWD-----GPISPIWTLDPESPGS-----TITTSNAPEK	77	:	:
QY	138	QAAGEY-LLIHQSERANYTVLFVNVRDYLQY-----LRRPYFRKMENODALI-CIS	187	:	:
Ddb	78	NTGYRCTELEPDMAGSTTI-----HLVYKDPAHSNNLLAQEVTVVGEQEAFLPCL-	128	:	:
QY	188	EGVPETVEMVLCSHRESCKEGBPAVRKRKEKVLUHELFGTDIR-----CCARNA	237	:	:
Ddb	129	--ITDPALK-----DSVSLMRGGROVLRAKTYFFSPWRGSIIIRAKAVLDNSTYVCKTMV	181	:	:
QY	238	LGRECTKL-FTIDLN---OAQOSTL-POLFUKV-CEPLWIRCKAIHVNHGFLGTWELED	290	:	:
Ddb	182	NGRSTSTGIWLKNRVHPPEPOIKLEPSKLVRIEAAAQIVCSATNAEVENVIILRGD	241	:	:
QY	291	KALE--EGSYFEMSTYSTNRTMIRILLAPVSVGNRDYGYTC--SSSKHPSQSALYT--I	345	:	:
Ddb	242	TKLEIPLNSDFQDNYYKYKRAL-----SLNAVFDQAGIYSVCVASNDVGYTRTATNFQV	295	:	:
QY	346	LKGFINATSSQE-EYEIDPYEKFSVFREKAYPRIR-CTWIFSOASFPCEORGLE----	399	:	:
Ddb	296	VESAYLNLTSSEQLSLOEVSVGSDSLILTVHADADYPSIQHYNMWLYLGPFFE-DORKLEFI	354	:	:
QY	400	---DGYSISKFCDH--KNKPGEYIFAENDDAQTFKMFLLNRKKRPQVLANASQASQ---	452	:	:
Ddb	355	RAIYRYTFKLFILNRVKASBAGOYFLMAQNKAQGNNUITFELTRYPEVSVTWMPVNGSDV	414	:	:
QY	453	--CSSDGYPLPSWTWKCKSDKPNCTEEIPEGVWN---KKANRVFGOWSVSSSTLNMSSE	506	:	:
Ddb	415	LFCDVSGYPQPSVTWMCECRGHTDRCDQAALHLANDTHPEVLSOKPDFKVIIOSQLPIGP	474	:	:
QY	507	AKGLLVKCAYNMGTSCTETFLNSPGPPPTQDNISFYATIGLC---LPFIVLIVLIVL	563	:	:
Ddb	475	LKHNMITYFCKTHNSVGNSSQYFRAVSLGOSKOLPDSESLFTPVVVACMSVMSLLVLLLL	534	:	:
QY	564	CHKYKQFVESOLOMIQ-VTGPLDNEYFYVDPRDYEDLKWFEPRENLEFEKGVLGSGAF	622	:	:
Ddb	535	LYKIYKPKYQVWKIERYEG---NSYTFIDPTQLPYNEKMEFFRNINLOFGKTLCGAGAF	591	:	:
QY	623	GRVMATAYGISTGYTSIQVAVKMLKEKADSCEKALMSSELKMTMHLGHHDNIENVLLGC	682	:	:
Ddb	592	GKVEATEAFLGKGEDAVLKAVAKMLASTAHADEKALMSSELKIMSHLGQHENIVNLLGAC	651	:	:
QY	683	TLGSPVYLIFYECYCGDLLNYLRSKREKPHRTWTEIFKHNPFSSYPTFOAHSSNSSMPGSR	742	:	:
Ddb	652	THGGPVLVTEYCYYGDLNFLRKAEAMHGP-----SLSPGDSEGSDGSYKNIH	701	:	:
QY	743	EVLHHPPLDOLSGFNNGSIHSEDEIEYENOKRLAEBEEDLN-----VLTFFEDLLCFAYQ	797	:	:
Ddb	702	LEKKYVVRD---SGFSOGVDVTVEMRPVSTSSDSFPFKDQDKHEHRPBDELWDLHFSSQ	759	:	:
QY	798	VAKGMFEFLPKSCVHRDLAARNVLVTHGVKVICDFGLARDILDSYVVYVVRGNARLPVKW	857	:	:
Ddb	760	VAOGMAFLASKNCIHRDVARNVLLTSGHVAKIGDGFGLARDINNDSYVVKVGA-NLPVKW	818	:	:
QY	858	MAPESLIFEGIYTTKSDWSYVIGILLWEIFSUGVNPYPCIPVDANFYKLIQSGFKMEQPFYA	917	:	:
Ddb	819	MAPESIFDCVITVQSDWSYVIGILLWEIFSLGLNFPYPGIYHNNKFYKLVKDGYSMAQPVFA	878	:	:
QY	918	TEGIYFMQSCWAFDSKRPSPFNLTSLF	946	:	:
Ddb	879	PKNIIYSIMQSCWDLPTRRPTFOOICFLL	907	:	:
RESULT 13					
TVHUKT					
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - human N;Alternate names: mast/system cell growth factor receptor; tyrosine kinase					

C:Species: Homo sapiens (man)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 21-Jul-2000
C:Accession: S01426; PC1015; B41815; B41815; I37948; I56954; I54336
R:Yarden, Y.; Kuang, W.J.; Yang-Feng, T.; Coussens, L.; Mumentz, S.; Dull, T.J.; Ch-
EMBO J. 6, 3341-3351, 1987
A:Title: Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for
A:Reference number: S01426; MUID:8811521; PMID:2448137
A:Accession: S01426
A:Molecule type: mRNA
A:Residues: 1-976 <VAR>
A:Cross-references: GB:X06182; NID:g34084; PIDN:CAA29548.1; PID:g34085
R:Hu, W.X.; Cornu, F.; Andre, C.; Galibert, F.
Chinese Biochem. J. 7, 618-629, 1991
A:Title: Nucleotide sequence of two neighbouring fragments of human c-kit proto-oncogene
A:Reference number: PC1015
A:Accession: PC1015
A:Molecule type: DNA
A:Residues: 412-713 <HUV>
A:Note: article in Chinese with English abstract
R:Spritz, R.A.; Giebel, L.B.; Holmes, S.A.
Am. J. Hum. Genet. 50, 261-269, 1992
A:Title: Dominant negative and loss of function mutations of the c-kit (mast/stem cell
A:Reference number: A41815; MUID:92133600; PMID:1370874
A:Accession: A41815
A:Molecule type: DNA
A:Residues: 579-583; 'L', 585-589 <SPR>
A:Cross-references: GB:S78839; NID:g244084; PIDN:AAB21234.1; PID:g244085
A:Note: sequence extracted from NCBI backbone (NCBIN:78839, NCBIP:78842)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: B41815
A:Molecule type: DNA
A:Residues: 637-641; 'SPELPW' <SP2>
A:Cross-references: GB:S78843; NID:g244086; PIDN:AAB21235.1; PID:g244087
A:Note: sequence extracted from NCBI backbone (NCBIN:78843, NCBIP:78844)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: C41815
A:Molecule type: DNA
A:Residues: 556-560; 'GGDKWK' <SP3>
A:Cross-references: GB:S78845; NID:g244088; PIDN:AAB21236.1; PID:g244089
A:Note: sequence extracted from NCBI backbone (NCBIN:78845, NCBIP:78846)
A:Note: disease-related mutant from patient with piebaldism
R:Giebel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.
Oncogene 7, 2207-2217, 1992
A:Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth
A:Reference number: I37948; MUID:93064697; PMID:1279499
A:Accession: I37948
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-976 <RES>
A:Cross-references: EMBL:X69301; NID:g34089; PIDN:CAA49159.1; PID:9825686
A:Note: an alternative splice form omitting residues 510-513 is described
R:Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.
Jpn. J. Cancer Res. 84, 1136-1144, 1993
A:Title: Characterization of the promoter region of the human c-kit proto-oncogene.
A:Reference number: I56954; MUID:94103107; PMID:7506248
A:Accession: I56954
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:S67773; NID:g459358; PIDN:AAB29529.1; PID:g459359
R:Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukui, K.
Hum. Mol. Genet. 2, 1499-1500, 1993
A:Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) prot
A:Reference number: I54336; MUID:94061059; PMID:7694728
A:Accession: I54336
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 242-250 <RE3>
A:Cross-references: GB:S67686; NID:g460545; PIDN:AAD13996.1; PID:g4261696
C:Genetics:
A:Gene: GDB:KIT
A:Cross-references: GDB:I120117; OMIM:164920
A:Map position: 4q12-q12

A: Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 6
 A: Note: defects in this gene may result in piebaldism
 C: Function:
 A: Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C: Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C: Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related
 rotein; tyrosine-specific protein kinase
 F: 1-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <MAT
 F: 1-509/514-976/Product: protein-tyrosine kinase kit precursor, short form #status predi
 F: 1-22/Domain: signal sequence #status predicted <SIG>
 F: 23-976/Product: protein-tyrosine kinase kit #status predicted <MAT>
 F: 23-520/Domain: extracellular #status predicted <EXT>
 F: 51-99/Domain: immunoglobulin homology <IMM1>
 F: 129-188/Domain: immunoglobulin homology <IMM2>
 F: 226-292/Domain: immunoglobulin homology <IMM3>
 F: 328-394/Domain: immunoglobulin homology <IMM4>
 F: 423-493/Domain: immunoglobulin homology <IMM5>
 F: 521-543/Domain: transmembrane #status predicted <TM>
 F: 544-976/Domain: intracellular #status predicted <INT>
 F: 587-931/Domain: protein kinase homology <KIN>
 F: 595-603/Region: protein kinase ATP-binding motif
 F: 58-97, 136-186, 233-290, 428-491/Disulfide bonds: #status predicted
 F: 130, 145, 283, 293, 300, 320, 352, 367, 463, 486/Binding site: carbohydrate (Asn) (covalent) #s
 F: 623, 640, 792/Active site: Lys, Glu, Asp #status predicted
 F: 797, 810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.7%; Score 1196; DB 1; Length 976;
 Best Local Similarity 31.4%; Pred. No. 6,8e-54;
 Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;

QY 36 CVLISHENNGSACKPSSYRMVGRSPEDLOCTPRRQSEGTVEATVVAESGSITLQVQ 95
 DB 12 CVLLLLRVQSGSQPS-----VSPGSPSPSIHPGKSDLIIVRGD-EIRLL 57
 QY 96 LATPGDLSCLVFKHSSLGCPHFQDLQNRGIVSMALLNVTQAGEYLLHQSERANTV 155
 DB 58 CTDPGFKV--WTFE---ILDETENKQEWITEKA-----EATNTGKYCTNKHGLSNIY 108
 QY 156 LFTVNRD-TOLYVLRPRFKMENQDALLCI-----SEG-----VPE 192
 DB 109 VF---VRDPAKFLVDRSLYCKEDNTLVRCPLDPEVNTSLKCCQCKPLPKDLRFPD 165
 QY 193 PTVEWVLCSSHRE-----SKREGEPAVVRKEEVLHFLGTDIRCCARNALGRECTK 244
 DB 166 PKAGIMIKSVKRAYHRLCHCSVDQEG-----KSVLSEKFLKVRPAFK-----A 210
 QY 245 LFTIDLNOAQOSTLPQLFLKVGEPFLWIRCKAIHNVHNGFGLTWELEDK--ALEE----- 295
 DB 211 VPVVSVSKA-----SYLLREGEFEFTVTCTIKDVSSSYSTWKRENSQTKLQEKYNSWH 264
 QY 296 GSYPEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILE---KGFN 352
 DB 265 GDF-----NYEQATLT-----ISSARVNDGVMFCYANNTFGSANVTTTLEVDKGFN 314
 QY 353 A-TSSQEEYIDPYKFCFSYRFRAYPR-IRCTWIFSOASF--PCEQRGLEGGYSISKFC 408
 DB 315 IFPMINTVFVNDGENVDLIVVEYAEKPEHQOQIYNNRTDTEKWDYPKSENEINRYV 374
 QY 409 DHKN-----KPEYIFYAENDDAQFTKMTLIRKKPQVLA--NASASQSCSSDGY 459
 DB 375 SELTLRLKGTETGTYTFLVNSNDVNAIAFNVYVNTKPEITLYDRLVNGMLQCVAAGFP 434
 QY 460 LPSWTWKCKSDKSPNCTEEI-PEGVNNKANRKYVFGQVSSSTLNMSEAGKLLVKCCAY 518
 DB 435 EPTIDWTFCTGTEORCSASVLPVDQTLNNSGPPFGKLVVOSSIDSSAFKINGTVECKAY 494
 QY 519 NSMGTCTETFLNSPGPPPTIQDN-----ISFYATIGLCPLFIVVLIVLIC 564
 DB 495 NDVGKT--SAYFN-----FAFGKNKEQIHPHTLTPLLIGFVIVAGM-----MCIIVMLT 544
 QY 565 HKYKKQFYESQLQMI-QVTGPLDNEYFYVDRDYEDLKWEPFRENLEFGKVLGSGAFG 623
 DB 545 YKYLQPMYEVQWKVVEING---NNVYIDPTQLPYDHKWEFPNRNLSFGKTLGAGAFG 601

QY 624 RVMNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSELKMMTHLGHHDNIYNLLGACT 683
 DB 602 KVEATEAYGLIKSDAAMTVAVKMLKSAHLTEREALMSELKVLISYLGNNHNINLLGACT 661
 QY 684 LSGPVILFEYCCYGDLLNLKSRKRFHRTWTEIFKEHNFSSYPTFOAHNSMSPGSR 743
 DB 662 IGGPTLVITEYCCYGDLLNFKRRKDSFCSKQEDHAEAAALYNLLHSSKSSCSDSTNEY 721
 QY 744 VOLHPPLDOL---SGFNGNSIHSEDIENOKRLAEEDDLNLTFFEDLLCFAYQVAK 800
 DB 722 MDMPKGVSVVPTKADKRKSRVIGSYIERDVTIPAIMEDEDELALDL---EDLLSFSYQVAK 778
 QY 801 GMEFEFKSCVHRDLAARNVLTGVKVKICDFGLARDILSDSSVYVYVGNARLPVKWMAP 860
 DB 779 GNAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVYVGNARLPVKWMAP 838
 QY 861 ESLFEGIYTKSDVMSYIGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATEG 920
 DB 839 ESIFNCVYTFSDVMSYIGIFLWELFSLGSSPYGMPVDSKFYKMKIEGFRMLSPHEAPAE 898
 QY 921 IYFVQSCWAFDSRRKRPSPFNLTSLFGLCOLAEA 953
 DB 899 MYDIMKTCWDADPLKRPTEKQIVQLIEKQISES 931

RESULT 14
 T30816
 macrophage colony-stimulating factor receptor - Japanese pufferfish
 C: Species: Fugu rubripes (Japanese pufferfish)
 C: Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
 C: Accession: T30816
 R: How, G.F.; Venkatesh, B.; Brenner, S.
 Genome Res. 6, 1185-1191, 1996
 A: Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for
 A: Reference number: 220882; MUID: 97129405; PMID: 8973913
 A: Accession: T30816
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-975 <HW>
 A: Cross-references: EMBL:U63926; NID: g1752706; PID: g1752708; PIDN: AAC60063.1
 C: Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

Query Match 22.6%; Score 1188.5; DB 2; Length 975;
 Best Local Similarity 33.0%; Pred. No. 1.6e-53;
 Matches 298; Conservative 153; Mismatches 291; Indels 161; Gaps 25;

QY 164 TOLYVLRPRYFRKMNQDALLCISEGVPPTVWVLCSSHRESCKEEGPAVVRKEEKVLH 223
 DB 124 TSLRVVRK-----EGEDYLLPCLLTDPDPAATDLGLRMDNGTTVPPEMNTYVYRHRGLIR 177
 QY 224 EL---FGTDIRCCARNALGRECTKLTIDLNOAQOSTLPQLFLK-----VGEPLWIRC 273
 DB 178 SIQPSFNADYVCTAKVGVEKTSKTSFINVIQKLRRP-PYVLEMDYEVYRIVGEEQLQIRC 236
 QY 274 KAIHNVHNGFGLTWELEDKA---LEGSIYEMSTYSTNRTMIRILLAFVSSVGNDRGTGYT 330
 DB 237 MTHNPNFNNTWNTYTKSRVTIEE---RVRSSGENRLDIQSILT-ISAVDLADTGNIS 291
 QY 331 CSSSKHP---SOSALVTILEKGF-----NATSSQEEYEDIDPYKFCFSYRFRAYPR 379
 DB 292 CIGTNEAGVNSNTYLLVVEKPYIRLWNPOLIPKLASQGLSVENEGEDLELGVWVEAYPQ 351
 QY 380 IRCTWIFSOAFPCQDRGLEGGYSISKFCDHK-----TDRWHTPTSPSTSMQEHYHARLQKRMNAQ 412
 DB 352 I-----TDRWHTPTSPSTSMQEHYHARLQKRMNAQ 384
 QY 413 KPEYIFYAENDDAQFTKMTLIRKKP-QVLNANASASQSCSSDGYPLPSTWTKKCSDK 471
 DB 385 EQGQYTFYAKSNLANGSISFHKYMKQPIAVVRWENITTLTCTSEGYPAQIIWYQCSGI 444
 QY 472 SPNCT-----BEIPEGVNNKANRKYVFGQVSSSTLNMSEAGKLLVKCCAYNSMGT 523

```
Db 445 RPTCNGNNTGLPKONHPOAL-TVEQREBEYGAIVESVFTVGLSNHRMTVECAVFNLCV 503
Qy 524 SCETIFLSPGPPFIQNI SFYATIGLC--LPFTIVLVILVICHKKOFRYESOLOMIO 581
Db 504 SSDTFTVE-----VSKLFTSTLIGAGVLAIFELLVFLYIKOKPREIRWKIIE 556
Qy 582 VTGPLDNEYFYVDFRDEYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISK-TGVSI 640
Db 557 ARE--GNNYTFIDPTQLPYNEKWEFPKGLKGLGAGAFKGVVEATAFGLGEDKNTL 614
Qy 641 QVAVKMLKEKADSCKEALMSLKMTHLGHHDNIVNLLGACTLSGPPYLIFFEYCCYGLD 700
Db 615 RVAVKMLKANAHSDREALMSLKLSHLGHHDNIVNLLGACTYGGPVLVITEYCSLGLD 674
Qy 701 LNYLSKSKKEFHRTWTEIFKEHNFSYPTFOAHN-----SSMPGR 742
Db 675 LNFLKQKATFVNLVNI-----PEIMENSNDYKNICNOKWIRSDSGISSTSS 725
Qy 743 EYQLHPPDLQDSGFNGNSIHSDEIEYENOKRLABEEEDNLVLTFFEDLLCFAYQVAKGM 802
Db 726 YLEMRP-----SQOSHTEASGRKSLCEDNGD--WPLDIDDLRLRESLQVAQGL 770
Qy 803 EPLEFKSVHRDLAARNVLVTGKVKKICDFGLARDILSDSSYVVRGNARLPVKWMAPE 862
Db 771 DFLASRNCIHRDVAARNVLLTDKRAKICDFGLARDIMNDSYVVKGNARLPVKWMAPE 830
Qy 863 LPEGIYTKSDVMSYGILLWEIFSLGVNPGIPVDANFYKLIQSGFKMEOPFYATEGIY 922
Db 831 IFDCVYTVOSDVMSYGILLWEIFSLGVNPGIPVDANFYKLIQSGFKMEOPFYATEGIY 890
Qy 923 FYMQSCWAFDSRKRPFPNLTSLFGLCOLA--EAEACIRTSIHLPKQAAPOORGGI 980
Db 891 MIMKWCWNLPEPTFTFSMISOMINRLGQDOEKLIYRNVO-PEQVAE----GEACDE 945
Qy 981 POR 983
Db 946 PKR 948

RESULT 15
JN0677
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken
N:Alternate names: tyrosine kinase receptor kit
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JN0677
R:Sasaki, E.; Okamura, H.; Chikamura, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.
Gene 128, 257-261, 1993
A:Title: Cloning and expression of the chicken c-kit proto-oncogene.
A:Reference number: JN0677; PMID:9329295; PMID:7685729
A:Accession: JN0677
A:Molecule type: mRNA
A:Residues: 1-960 <SAS>
A:Cross-references: DDBJ:U13225; NID:g303532; PIDN:BAA02506.1; PID:g303533
A:Experimental source: brain
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
rotein kinase
F:1/24/Domain: signal sequence #status predicted <SIG>
F:25-960/Product: tyrosine kinase receptor #status predicted <MAT>
F:314-380/Domain: immunoglobulin homology <IMM>
F:573-916/Domain: protein kinase homology <KIN>
F:581-589/Region: protein kinase ATP-binding motif
F:76,135,149,269,286,306,318,338,356,453,469/Blinding site: carbohydrate (Asn) (covalent)

Query Match 22.4%; Score 1181; DB 1; Length 960;
Best Local Similarity 31.8%; Pred. No. 3.9e-53;
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;

Qy 45 GSSAGKPSYRMVRGSPEDLOCTPRROSEGTVEAAVVEAEGSITLQVOLATPGLDSC 104
Db 24 GSPHEESSLVNKGELRLKCN-----EEGPTVWTFNQSDPSAKTR 65
```

Search completed: May 27, 2003, 14:35:21
Job time : 32.4866 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:49 ; Search time 17.9909 seconds

(without alignments)
2286.959 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLLVLSV.....RGLLRAQSPQVRQVTHRSR 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5264	100.0	992	1 FLT3_MOUSE	Q00342 mus musculus
2	4429.5	84.1	992	1 FLT3_HUMAN	P36888 homo sapien
3	1258.5	23.9	980	1 KEMS_FELCA	P13369 felis silve
4	1249.5	23.7	978	1 KEMS_FSVMD	P00545 feline sarc
5	1223	23.2	975	1 KIT_MOUSE	P05532 mus musculus
6	1222.5	23.2	978	1 KIT_CAPHI	Q28317 capra hircu
7	1221.5	23.2	972	1 KEMS_HUMAN	P07333 homo sapien
8	1219	23.2	977	1 KEMS_MOUSE	P09581 mus musculus
9	1215	23.1	977	1 KIT_BOVIN	P43481 bos taurus
10	1210	23.0	978	1 KEMS_RAT	Q00495 rattus norv
11	1207	22.9	975	1 KIT_CANFA	O97799 canis famil
12	1196	22.7	976	1 KIT_HUMAN	P10721 homo sapien
13	1181	22.4	960	1 KIT_CHICK	Q08156 gallus gall
14	1178.5	22.4	978	1 KIT_FELCA	Q28889 felis silve
15	1157.5	22.0	1088	1 PGDS_RAT	P20786 rattus norv
16	1152.5	21.9	1089	1 PGDS_HUMAN	P16234 mus musculus
17	1142.5	21.7	1089	1 PGDS_MOUSE	P26618 mus musculus
18	1132	21.5	1087	1 PGDS_XENLA	P26619 xenopus lae
19	1098	20.9	1098	1 PGDR_MOUSE	P05622 mus musculus
20	1078	20.5	1106	1 PGDR_HUMAN	P09619 homo sapien
21	988.5	19.0	1338	1 VGR1_HUMAN	P17948 h vasculat
22	984.5	18.7	1336	1 VGR1_MOUSE	P53767 rattus norv
23	983.5	18.7	1333	1 VGR1_MOUSE	P35969 mus musculus
24	967.5	18.4	370	1 KIT_FSVHZ	P04048 feline sarc
25	965	18.3	1348	1 VGR2_COTJA	P52583 coturnix co
26	961.5	18.3	1356	1 VGR2_HUMAN	P35968 homo sapien
27	954.5	18.1	1298	1 VGR3_HUMAN	P35916 homo sapien
28	952	18.1	1363	1 VGR3_MOUSE	P35917 mus musculus
29	934	17.7	1343	1 VGR2_RAT	O08775 rattus norv
30	926.5	17.6	1367	1 VGR2_MOUSE	P35918 mus musculus
31	818	15.5	813	1 FGR2_XENLA	Q03364 xenopus lae
32	810.5	15.4	823	1 CBK3_CHICK	P18461 gallus gall
33	791	15.0	821	1 FGR2_HUMAN	P21802 homo sapien

34	789	15.0	806	1 CEK2_CHICK	P18460 gallus gall
35	780	14.8	821	1 FGR2_MOUSE	P21803 mus musculus
36	766.5	14.6	819	1 FGR1_CHICK	P21804 gallus gall
37	761.5	14.5	654	1 BFR2_HUMAN	Q01742 homo sapien
38	758.5	14.4	812	1 FGR1_XENLA	P22182 xenopus lae
39	754.5	14.3	822	1 FGR1_HUMAN	P11362 homo sapien
40	751.5	14.3	806	1 FGR3_HUMAN	P22607 homo sapien
41	751.5	14.3	822	1 FGR1_MOUSE	P16092 mus musculus
42	748.5	14.2	822	1 FGR1_RAT	Q04589 rattus norv
43	732	13.9	801	1 FGR3_MOUSE	Q61851 mus musculus
44	698	13.3	802	1 FGR4_HUMAN	P22455 homo sapien
45	697	13.2	808	1 FGR4_MOUSE	Q03142 mus musculus

ALIGNMENTS

```

RESULT 1
FLT3_MOUSE
ID FLT3_MOUSE STANDARD; PRT; 992 AA.
AC Q00342;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase
DE receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).
GN FLT3 OR FLT-3 OR FLK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91292518; PubMed=1648448;
RA Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;
RT "A receptor tyrosine kinase specific to hematopoietic stem and
RT progenitor cell-enriched populations.";
RL Cell 65:1143-1152(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92019834; PubMed=1656368;
RA Rosnet O., Marchetto S., Delapeyriere O., Birnbaum D.;
RT "Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the
RT PDGFR/Csflr family.";
RL Oncogene 6:1641-1650(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93205405; PubMed=8384358;
RA Maroc N., Rottapel R., Rosnet O., Marchetto S., Lavezzi C.,
RA Mannoni P., Birnbaum D., Dubreuil P.;
RT "Biochemical characterization and analysis of the transforming
RT potential of the FLT3/FLK2 receptor tyrosine kinase.";
RL Oncogene 8:909-918(1993).
CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-
CC ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL; M64689; AAA37634.1; -
 DR EMBL; X59398; CAA42041.1; -
 DR PIR; A39931; A39931.
 DR HSP; P11362; IFGK.
 DR MGD; MGI:95559; Fik3.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR001824; RTK_kinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00410; IG_Like; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 992 FL CYTOKINE RECEPTOR.
 FT DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 545 564 POTENTIAL.
 FT DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 611 946 PROTEIN KINASE.
 FT NP_BIND 617 625 ATP (BY SIMILARITY).
 FT BINDING 645 645 ATP (BY SIMILARITY).
 FT ACT_SITE 814 814 BY SIMILARITY.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 150 150 R -> A (IN REF. 2).
 FT CONFLICT 242 242 C -> S (IN REF. 2).
 FT CONFLICT 726 726 S -> F (IN REF. 2).
 FT CONFLICT 957 979 IYONRRLPSREAGSEPP (IN REF. 2).
 FT CONFLICT 983 983 R -> A (IN REF. 2).
 SQ SEQUENCE 992 AA; 112639 MW; 407A087853372100 CRC64;

Query Match 100.0%; Score 5264; DB 1; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMRVGRS 60
 DB 1 MRALQSRDRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSRYMRVGRS 60
 QY 61 PEDLOCTPRROSECTVVEAATVEAESITLQVOLATPGDLSCLWFKHSLGCOPIHD 120
 DB 61 PEDLOCTPRROSECTVVEAATVEAESITLQVOLATPGDLSCLWFKHSLGCOPIHD 120
 QY 121 LQNRGIVSMALLNTVTOAGEYLLHIOSEANRYTVLFTVNVYRDYQLYVLRPRYFRKMENQ 180
 DB 121 LQNRGIVSMALLNTVTOAGEYLLHIOSEANRYTVLFTVNVYRDYQLYVLRPRYFRKMENQ 180
 QY 181 DALLCISEGPEPTVEVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCCARNALGR 240
 DB 181 DALLCISEGPEPTVEVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCCARNALGR 240
 QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPDLWTRCAIHVNHGFGITWELEDKALEGSGYFE 300
 DB 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPDLWTRCAIHVNHGFGITWELEDKALEGSGYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPQSALVTILEKGFINATSSQEEY 360

DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPQSALVTILEKGFINATSSQEEY 360
 QY 361 EIDPYEKFCFSVRKAYPRICTWIFSOAPCEORGLDGYISKFCDHKNKPGEYIFY 420
 DB 361 EIDPYEKFCFSVRKAYPRICTWIFSOAPCEORGLDGYISKFCDHKNKPGEYIFY 420
 QY 421 AENDDAOFTKMTLNIRKKPOVLANASQASQSSDGYPLPSWTWKKCKSDKSPNCTEELP 480
 DB 421 AENDDAOFTKMTLNIRKKPOVLANASQASQSSDGYPLPSWTWKKCKSDKSPNCTEELP 480
 QY 481 EGVNKKANRKFVGOVSSSTLNWSEAGKLLVKCCAYNSMGTSCETIFLNSPGGPEPFIQ 540
 DB 481 EGVNKKANRKFVGOVSSSTLNWSEAGKLLVKCCAYNSMGTSCETIFLNSPGGPEPFIQ 540
 QY 541 DNISFYATIGLCLPFIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEFYFVDFRDY 600
 DB 541 DNISFYATIGLCLPFIVVLIVLICHYKQFRYESQLOMIQVTPGLDNEFYFVDFRDY 600
 QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660
 DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660
 QY 661 SELKMMTHLGHHDNIIVNLGACTLSGPVYLFEYCCYGDLLNLYLSKREKFRHTWTEIEFK 720
 DB 661 SELKMMTHLGHHDNIIVNLGACTLSGPVYLFEYCCYGDLLNLYLSKREKFRHTWTEIEFK 720
 QY 721 EHNFSSTYPTFOAHNSNMPGSGREVOLHPDLQSLGPNFNGSIHSEDEIEYENOKRLAESEE 780
 DB 721 EHNFSSTYPTFOAHNSNMPGSGREVOLHPDLQSLGPNFNGSIHSEDEIEYENOKRLAESEE 780
 QY 781 EDLNVLTPEDDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIL 840
 DB 781 EDLNVLTPEDDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIL 840
 QY 841 SDSSVYVRGNARLPVKWMAPELSFEGITYIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
 DB 841 SDSSVYVRGNARLPVKWMAPELSFEGITYIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
 QY 901 FYKLQSGFKMEOPPYATEGIVFMQSCWAFDSRSPNPNTSLFGCOLAAEEACIRT 960
 DB 901 FYKLQSGFKMEOPPYATEGIVFMQSCWAFDSRSPNPNTSLFGCOLAAEEACIRT 960
 QY 961 SIHLPKQAAPQQRGLRAQSPQORVKIHRERS 992
 DB 961 SIHLPKQAAPQQRGLRAQSPQORVKIHRERS 992

RESULT 2
 ID FLT3_HUMAN STANDARD; PRT: 993 AA.
 AC P36888; O13414;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).
 GN FLT3 OR STK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94119906; PubMed=7507245;
 RA Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P.,
 RA Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.;
 RA "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in
 RT CD34+ human bone marrow cells and is involved in the proliferation of
 RT early progenitor/Stem cells".
 RL Proc. Natl. Acad. Sci. U.S.A. 91:459-463(1994).
 RN [2]
 RP SEQUENCE FROM N.A.

Db 838 SDSNVTVRGNARLPVKWMAPESEFEGYTIKSDVWSYGLLWEIFSLGYNVPYGPVVDAN 897
Qy 901 FYKLOSFKKPEPYATGYFYVMSQWAFDSRRSPFNLTSLGCOLAEAEAC--- 957
Db 898 FYKLLQNGFKMDQPPYATGYFYVMSQWAFDSRRSPFNLTSLGCOLADREAEAYON 957
Qy 958 ---IRTSIHLPKQAAPOQRG-GLRAQSPQRQVK 986
Db 958 VDRGVSECPHTYQNRPFPSRENDLGLLSQAQV 991
RESULT 3
ID KMS_FELCA STANDARD; PRT: 980 AA.
AC P13369;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR FMS.
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89077553; PubMed=2849512;
RA Woolf J., McAuliffe A., Rohrschneider L.R.;
RT "Activation of the feline c-fms proto-oncogene: multiple alterations
are required to generate a fully transformed phenotype.";
RL Cell 53:965-977(1988).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J03149; AAA30811.1;
CC FIR; A31636; YVCTMD.
CC HSP; P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; ig; 3.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00410; IG-like; 4.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1.
FT 19
FT POTENTIAL.

FT CHAIN 20 980
FT DOMAIN 20
FT TRANSMEM 510
FT DOMAIN 536
FT DOMAIN 24
FT DOMAIN 107
FT DOMAIN 204
FT DOMAIN 299
FT DOMAIN 398
FT DOMAIN 579
FT NP_BIND 585
FT BINDING 613
FT ACT_SITE 776
FT DISULFID 42
FT DISULFID 127
FT DISULFID 224
FT DISULFID 417
FT MOD_RES 807
FT CARBOHYD 45
FT CARBOHYD 73
FT CARBOHYD 94
FT CARBOHYD 153
FT CARBOHYD 275
FT CARBOHYD 286
FT CARBOHYD 302
FT CARBOHYD 335
FT CARBOHYD 410
FT CARBOHYD 477
FT CARBOHYD 490
SQ SEQUENCE 980 AA; 108506 MW; 45CF661E97CF6FF CRC64;
Query Match 23.9%; Score 1258.5; DB 1; Length 980;
Best Local Similarity 32.7%; Pred. No. 9.8e-77;
Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;
Qy 75 TVYEATVEVAESGISTLOVLATPGDLISCL-----W---VFKHSSLGCCOPHFQDNRG 125
Db 13 TAWHAQGVPIQPSGPELVVEPGTIVTLRCVNGSVEMDGPISPHWNLDLDPSSI----- 68
Qy 126 IVSMALINVTETQAGEYLLHI-----QSERANYTVLFTVNRDQLYVL---RRPY----- 173
Db 69 ---LTTNNATFQNTGY--HCTEPNPGGNATI-----HLVVKDPAWPVKVLAQE 114
Qy 174 FRKMENODALL-CISEGVPEPTVEWLCSSHRECKEGPVAVRKEKVLHFLFTDI-- 230
Db 115 VTVLEGDALLPCL---LTDPALE-----AGVSLVRVRGPRVLRQTNYSFSPWHGFTIHK 166
Qy 231 -----RCCARNALGRECTKL-----FTIDLNOAQOSTL-PQLFLKV-GEPLWIRCK 274
Db 167 AKFTIENHVQCSAR-VDGRVTVTSMGIWLVKQKDISGPAITLLEPAELVRIQGEAAQIVCS 225
Qy 275 AIHVNHGFLTWELEDKALEGYSFEMSTYNTNMTIRILLAFVSSVGRNDTGYTCSSS 334
Db 226 ASNIDVNFDSVLRHGDTKL---TISQSDFDHNRVQ-KVLTLNLDHVSFQDAGNYSCTAT 281
Qy 335 K---HPSQSALVTILEKGFINATSSQE-EYEDIPYKEKCFSVRFKAYPRIC-TWI---- 385
Db 282 NAWGNHSASMVFRVVSAYLNLTSQSLQELQVTVGKEVDLQVKVEAYPGLESFNNYTLGP 341
Qy 386 FSQASFPCEQRLGLEDGYSISKFD-----HKNKPGEYIFAENDDAQFTKMTFLNIRKKPQ 441
Db 342 FSDYQDKLDFVTIKDTYRTYTLTLPLRKRSEAGYSFLARNAGQGNALTFTLTUYRPE 401
Qy 442 VLNASASOAS-----CSSDGYPLPSTWTKCKSDKSPNCTEE---IPEGVNNKANRKFVG 494
Db 402 VRVTMTLINGSDTLLCEASGHPQPSVTWVOCRSHTRDCRDESAGLVLESHSEVLSQVPH 461
Qy 495 QWVSSSTLNNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFFITQDNISFYATIGLC--- 552
Db 462 EVIVHSLLAIGTLEHNRHYECRAFNVSNGSSQTFWPISGAHTQLPDELFTPVLLTCS 521
Qy 553 -LPFTIVLIVLICHYKKQFRIESQLOMIQVGTGLDNEYFYVDPRDYEDLAKWEFFREN 611

```
Db 522 IMALLLLLLLLLLLYKQKPKYQVRKKIIE--SYEGNSYTFIDPTQLPYNEKWEFFRNLL 579
QY 612 EFGKVLGSAFGVMATATYIGSKTGVSTQVAVKMLKEKADSCKEKALMSELKMMTHLGH 671
Db 580 QFGKTLGAGAFKGVVEATAGFLGKEDAVLKVAVKMLKSTAHADEKEALMSELKIMSHLGQ 639
QY 672 HONIVNLLGACLTSGPVYLIFECYCCYCDLLNLYLSKRE-----KF 711
Db 640 HENIVNLLGACHTGGPVVLVITEYCCYCGDLLNLFURROAEAMLGSLSVGOPEAGAYKNI 699
QY 712 HRTWTETFEKHNFPSTYPTFOAHSSNMPGSREVQLHPPLDOLSGNGNSIHSDEYEN 771
Db 700 HLEKKVVRDSDES-----SOGVDYTVEMRPVSTSSNSDSFSEDL---- 740
QY 772 OKRLAEEEDLNVLFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKIC 831
Db 741 -----GREDRPLELRDOLLHPSQVAGCMAPLASKNCITHRDVAARNVLLTSGRVAKIG 793
QY 832 DFLGARDILSDSYVYVGRNARLPVKWMAPESELEFGIYTIKSDWSYGILLWEIFSLGVNP 891
Db 794 DFLGARDIMDSNYIYKGNARLPVKWMAPESEIDFCYTVQSDWSYGILLWEIFSLGLNP 853
QY 892 YGIPVDANFYKLIQSGFKMEQFYATEGTYFVMQSCWAFDSKRPSPNLTSTFLGCOLA 951
Db 854 YGILVNSFYKLVKDCYQMAQAPAFAPKNIYSIMQACWALEPRTTRPTFOQICSL--QKQ 911
QY 952 EAEACIRTSIHLPKQAAPQ-----RGLRAQSPQRQ 984
Db 912 AQEDRRVPVNTLPPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 953

RESULT 4
KFMS_FSVMD
AC P00545; O86597; STANDARD; PRT; 978 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).
GN FMS.
OS Feline sarcoma virus (strain McDonough).
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=111778;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84119469; PubMed=6582485;
RA Hampe A., Gobet M., Sherr C.J., Galibert F.;
RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows
RT unexpected homology with oncogenes encoding tyrosine-specific protein
RT kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE=9201516; PubMed=1833563;
RA Smola U., Hennig D., Hadwiger-Fangmeier A., Schuetz B., Pfaff E.,
RA Niemann H., Tamura T.;
RT "Reassessment of the v-fms sequence: threonine phosphorylation of the
RT COOH-terminal domain.";
RL J. Virol. 65:6181-6187(1991).
CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY
CC STIMULATING FACTOR 1 (CSF-1).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS
CC POLYPEPTIDE.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: K01643; AAA3045.1; -
CC FIR: A00654; TVMVD.
CC HSP: P11362; IFGK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_c2.
CC InterPro: IPR003598; Ig_c2.
CC InterPro: IPR003600; Ig_like.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; ig; 3.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 2.
CC SMART: SM00410; IG_like; 4.
CC SMART: SM00408; IGC2; 1.
CC SMART: SM00219; TyrKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Polyprotein; Tyrosine-protein kinase; Oncogene; Transferase; Receptor;
CC Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;
CC Immunoglobulin domain; Repeat.
CC DOMAIN 1 543 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 544 568 POTENTIAL.
CC DOMAIN 569 978 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 58 138 IG-LIKE C2-TYPE DOMAIN 1.
CC DOMAIN 141 231 IG-LIKE C2-TYPE DOMAIN 2.
CC DOMAIN 238 332 IG-LIKE C2-TYPE DOMAIN 3.
CC DOMAIN 333 431 IG-LIKE C2-TYPE DOMAIN 4.
CC DOMAIN 432 536 IG-LIKE C2-TYPE DOMAIN 5.
CC DOMAIN 613 942 PROTEIN KINASE.
CC NP_BIND 619 627 ATP (BY SIMILARITY).
CC BINDING 647 647 ATP (BY SIMILARITY).
CC DISULFID 76 118 POTENTIAL.
CC DISULFID 161 211 POTENTIAL.
CC DISULFID 258 312 POTENTIAL.
CC DISULFID 451 516 POTENTIAL.
CC MOD_RES 841 841 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC MOD_RES 973 973 PHOSPHORYLATION.
CC CARBOHYD 79 79 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 107 107 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 128 128 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 309 309 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 320 320 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 336 336 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 369 369 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 511 511 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 524 524 N-LINKED (GLCNAC...) (POTENTIAL).
CC CONFLICT 714 714 L -> P (IN REF. 1).
CC CONFLICT 971 978 ORTPPVAR -> RGPPL (IN REF. 1).
CC SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;

Query Match 23.7%; Score 1249.5; DB 1; Length 978;
Best Local Similarity 33.2%; Pred. No. 3.9e-76;
Matches 328; Conservative 162; Mismatches 348; Indels 149; Gaps 32;

QY 75 TVTEATVEVAESGSIITQLVOLATPGDLSCI-----W---VFKHSLGCGPHFDLQNRG 125
Db 47 TAWHAQGVPIQPSGPELVPEPTVTLCVGVNGSVEMDGPISPHWNLDLPSSI----- 102
QY 126 IVSMALINVTETQAGEYLLHI-----QSERANYTVLFTVNVVDLTOLYVL--RRPY----- 173
Db 103 ---LTTNNATFQNTGTG---HCTEPGNPRGGNATI-----HLYVKDARPWKVLAQE 148
```

QY 174 FRKMENODALL-CISEGVPEPTVETVVLCSHRESCKEAGPAVVRKEEVHLHFGTDI-- 230
Db 149 VTVEGQDALLPCL---LTDPALE-----AGVSLVRVGRGPVLRQTNYSFSPHGHGTIHK 200
QY 231 -----RCCARNALGRECTKL-----FTIDLNAQOSTL-PQLFLKV-GEPLMIRCK 274
Db 201 AKFIENHVQCSAR-VDRGRTVTSMGILWKVKQDKISGPATLTLEPAELVRIQGEAAQIVCS 259
QY 275 AIVNHGFGLTWELEKALEEGSYFEMSYSTNRTMIRILLARVSSVGRNDTGYTCSSS 334
Db 260 ASNIDNVFVSLRHGDTKL---TISQOSDFHNRQY-KVLTLLNDLHVSFODAGNYSCTAT 315
QY 335 K---HPSQALVTLILEKGFINATSSQE-EYEIDPYEKFCFSVRKAYPRRC-TWI----- 385
Db 316 NAWGNHSAWVFRVESAYNLTSEGLLQEVTVGEKVDLOVKEAYPGLESEFNWYLG 375
QY 386 FSOASPCQORGLDGYISKFC-----HNKPGEYIFYAENDDAQFTKMTNIRKKPQ 441
Db 376 FSDYQDKLDFVTIKDYRYVTSTLSPLRKRSSEGRYSFLARNAGGQNALTFELTLRYPE 435
QY 442 VIANASASQAS-----CSSDGYPLPSWTWKCKSDKSPNCTEE---IPEGVWKNKANRVFG 494
Db 436 VRVTMTLINGSDFLLCEASGYPOPSVTWQCRSHTRDCDESAGLVLEDSHSEVLISQVPFY 495
QY 495 QWVSSSTLNMSAEGKLLVKCCAYNSMGTSCTETIFLNSPGPFPIODNIFATIGLC-- 552
Db 496 EVIVHSLAIGTLEHNTYECRAFNSVGNSSQTFWPISIGAHTPLDPELLFTPVLLTCMS 555
QY 553 -LPFIVVLIVLICKYKQPKRYESQLOMQVTGPDNEYFYVDFRDYEDYDLKWEFFREN 611
Db 556 IMALLLLLLLLLYKKQPKYQVRNKKIE--SVEGNSYTFIDPTQLPYNEKWEFFRNNL 613
QY 612 EFGKVLGSGAFGRVMATAYGISKTVSIOVAVKMLKEKADSCKEALMSKLKMMTHLGH 671
Db 614 QFGKTLTGAGFKVWEATAFGLGKEDAVLVKAVKMLKSTAHADKEALMSKLKIMSHLQ 673
QY 672 HDNIVNLGACTLSGPPVILFEYCCYGDLLNLYLSKRE-----KF 711
Db 674 HENIVNLGACTGGPVLVITECCYGDLLNLFRRQAEMLGPSLVGQDPEAGAGYKNI 733
QY 712 HRTWTEIFKEHNFSTPYTFOAHNSMSPGSRREVQLHPPLDQLSGFNGNSIHSDEIYEN 771
Db 734 HLEKVVYRRDSGFS-----SOGVDYVEMRPVSTSSNDGSFSEEDL----- 774
QY 772 OKRLAEEEDLNLVDFELLCFAYQVAKMEFLFCKSVHRDLAARNVLVTHGKVVKIC 831
Db 775 -----GKEDGRPLELDLHFSSQVQAQMAFLASKNCIHRDVAARNVLLTSGRVAKIG 827
QY 832 DFGLARDILSDSVVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNP 891
Db 828 DFGLARDIMDSNVIVAGNARLPVKWMAPESFDCVYTVQSDVWSYGILLWEIFSLGLNP 887
QY 892 YPGIPVDANFYKLQSGFKMEQPPYATEGIFYVQMSQWAFDSRRKRPFPNLTSLGCOLA 951
Db 888 YPGLVNSKFKYKLVKQGYQMAQAPAFKNIYSIMQACWALEPRTTRPTFQICSL--QKQ 945
QY 952 EABEACIRTSIHLPKQAA-----PQOR 973
Db 946 AQEDRRVPNTNLPSSSSSRLRPQOR 972

RESULT 5

KIT_MOUSE
ID KIT_MOUSE STANDARD; PRT; 975 AA.
AC P05532; Q61415; Q61416; Q61417;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/system cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT OR SL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=88296403; PubMed=2456920;
RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddie F.H.,
RA Besmer P.;
RT "Primary structure of c-kit: relationship with the CSF-1/PDGF
RT receptor kinase family -- oncogenic activation of v-kit involves
RT deletion of extracellular domain and C terminus.";
RL EMBO J. 7:1003-1011(1988).
RN [2]
RP SEQUENCE FROM N.A. (TRUNCATED FORM).
RC STRAIN=ICR;
RX MEDLINE=92331813; PubMed=1378413;
RA Rossi P., Marziani G., Albanesi C., Charlesworth A., Geremia R.,
RA Sorrentino V.;
RT "A novel c-kit transcript, potentially encoding a truncated receptor,
RT originates within a kit gene intron in mouse spermatids.";
RL Dev. Biol. 152:203-207(1992).
RN [3]
RP LIGAND.
RX MEDLINE=91006023; PubMed=1698611;
RA Tan J.C., Buck J., Levi E., Besmer P.;
RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a
RT fibroblast derived growth factor stimulates mast cells and erythroid
RT progenitors.";
RL EMBO J. 9:3287-3294(1990).
RN [4]
RP VARIANT W42 ASN-790.
RX MEDLINE=90100577; PubMed=1688471;
RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;
RT "The dominant W42 spotting phenotype results from a missense mutation
RT in the c-kit receptor kinase.";
RL Science 247:209-212(1990).
RN [5]
RP VARIANTS W37 LYS-582; W41 MET-660 AND W41 MET-831.
RX MEDLINE=90269214; PubMed=1693331;
RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,
RA Besmer P.;
RT "Molecular bases of dominant negative and loss of function mutations
RT at the murine c-kit/white spotting locus: W37, W41 and W41 and W41";
RL EMBO J. 9:1805-1813(1990).
CC -I- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DISEASE: WHITE-SPOTTING VARIANTS INDUCES SEVERE EFFECTS ON
CC PIGMENTATION, GAMETOGENESIS AND HEMATOPOIESIS. MICE HOMOZYGOUS
CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.
CC -I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -I- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00864; CAA68772.1; -
DR EMBL; X65997; CAA46798.1; -
DR EMBL; X65998; CAA46799.1; ALT_SEQ.
DR EMBL; X65998; CAA46800.1; -
DR PIR; S00474; TVMSKT.
DR HSP; P11362; IFGK.

Db 252 PQPHIAQVXKHNSWHRGDF-----NYERQEPLT-----ISSARVDSDGVMFCVANNTFGS 301
 Qy 340 SALVT---ILEKGFINATS-SQBEEYEDIDPEKEFCFSVRFKAYPR-IRCTWIFSQAFFPCE 394
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 302 ANVTTLKVVEKGFINISPVKNITVFVTCGENVDLVVEYEAEPKBEHQOVIYNRT -SA 359
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 395 ORGLE----DCYSISRKFDH-----KNKPGEYIFYAENDDAQAFMKFTLNRKKPKQVL 443
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 360 NKGRDYVKSNKNI RYVNQLRLTRLKGTGGTYTEFLVNSDSASASVTFNVVYNTKPEIL 419
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 444 A-NASASQASCSDGYPLPSWTWKKSCKSPNCTEEI-PEGVWNKANRKFVGQAVSSS 500
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 420 TYDLRINGMLQCVAEGPEPTIDWFCTGAEBQRCTTPVPSPDVQVQVSVSPGKLWVQS 479
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 501 TLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPCPPFIQDNISFYATIGLCGLPFIV-- 557
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 480 SIDSVFRHNGTVECKASNDVGKS--SAFEF-----PAFKEQIQAHLTFTPLLIGFVVAAG 533
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 558 ---VLIWLICHKKYKQPRYSQLOMI-QVTGPDLNEYFYVDRDREYDLKWEPFPRENLEF 613
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 534 AMGIIVMVLT VKYLQKPMYEVOQKVBEING---NNVYVIDPTQLPYDHKWEPFNRLSF 590
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 614 GKVLGSAGFRVMNATAYGISKTVCSIQAVAKMLKEKADCKEALMSCLKMMTHLGHHHD 673
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 591 KTGIGAAGFGKWEATAYGLIKSDAAHTAVAKMLKPSAHLTEREALMSCLKVLSYLGNHM 650
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 674 NIVNLGACTLSGPGVYLIEFYCVCGDLLNLRKRKF-----HRTWTEIFRKEHFSSY 727
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 651 NIVNLGACTVGGPTLVITEYCVCYGDLLNFLRRKRDSFIFSKOEQA EALYKNLLHSTE 710
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 728 PTOAHNSSM---PGSREVQLHPDL DOLSGFNNGNSTHSEDEIEYENQKRLABEEEEELN 784
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 711 PSCDS-SNEYMMDMPGVSYV-VPTKTOK-----RRSARIDSITYERTDVTAPAIMDDDELALD 763
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 785 VLTPEDILLCFAYQAVAKMEFELEKSCVHRDLAARNVLVTHGKVVVKICDFGLARDILSDSS 844
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 764 L---DDLTSFSYQAKAMAFSLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRNDNS 820
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 845 YVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNPYPGIPVDANFYKL 904
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 821 YVVGKNARLPVKWMAPESIFSCVYTTFESDVMSYGIFLWELFSLGSSPYPCMPVDSKFYKM 880
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Qy 905 IQSGFKMQEPPYATEGYIFYVWSQAWFDSRRKRSFPLNLSFLGCQLAEA 953
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 881 IKGFRMVSPBHAPAEMYDMYKTKCDWDAPLKRPTFKQVQVLIBEKQISDS 929
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
 KIT_CAPHI . STANDARD; PRG; 978 AA.
 AC Q28317;
 ID Q28317;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
 GN KIT.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 RN NCBI_TaxID=9925;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Shiba; TISSUE=Cerebellum;
 RX MEDLINE=97342548; PubMed=9199245;
 RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,
 RA Kitamura Y., Sawasaki T., Tachi C.;
 RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats
 RT and a novel alanine insertion encoding specific to goats and sheep in the
 RT kinase insert region.";
 RL Biochim. Biophys. Acta 1352:151-155(1997).
 CC !- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL

[illegible]

Df	I72	SQDYQSALMG-GR--KVMKSISIRLKVKVQIPGPPALVLPAELVRIGEEAAIVCSAS	227
Qy	277	HVNHGGLTWELEDKALEEGSFEMSTYSTRNRMIRILLAFVSSVGRNDTGXYTCSSS-	334
Df	228	SVDNFDFVLOHNTKL---AIPQSDFFHNRYQ-KVLTLNLDDQVDFQAGNSFCVASNV	283
Qy	335	--KHPSQSALVTILEGFINATSQE-EYEIDPYEFCEFSVRFKAYPRIR-CTWIFSQAS	390
Df	284	OGRH-STSMFRVRESAYLNLSEQNLIQEVTVGEGLNKVMVEAPYGLOGFNWTY----	338
Qy	391	FPCEQRLEGDGYISKPCDHKNP-----GEYIPIYAEN	423
Df	339	-----LGPFSDHQPEPKLANATTKTDTYRHTFTLSLPRLKSEAGRIYSFLARN	385
Qy	424	DDAQFTKMTLNLRKKPOV-----LANASQASCSDGVPLSPWTWKCKSCDKSPNCTEE	478
Df	386	PGWRALTFTLURYPEVSIVITWFINGSGT-LICAASGPQPNVTWLQCSGHTRDCEA	444
Qy	479	IPEGVWN----KANKRVFGQWSSSTLNNMSEAGKLLVKCCAYNSMGTSCETIFLNSPG	534
Df	445	QVLQWMDPPPEVLSQEPFHKVTVQSLTTVETLEHQTVECAHNSVGSGSWAFIPISAG	504
Qy	535	PFPFIQONISFYATIGLC----LPFVVILVICHKKYKQRFYESQLQMIOVTGPLNEFY	591
Df	505	AHTHPDEFLTPTVVACMSIMALLLLLLLLLKYKQPKYQVRWKIIE--SYEGNSYT	562
Qy	592	YVDPROYDLKWFEPRENFEGKLVGGAGFGRVMNATYGIKSTGVSIQAVYMKLKEKA	651
Df	563	FIDPTQLPYNKEWFEPPRNLFQKTLGAGAFGKVEATATGLKGEDAVLVAVYMKLSTA	622
Qy	652	DSCEKALMSLKMTHLHDHNTVLNLGACTLGGPVYLIFYECYCGDLLNYLRSREKF	711
Df	623	HADKEALMSLKMTHLGHQENIVNLGACTHGGPVLVITECYCYCGDLLNFLRRKAE--	680
Qy	712	HRTWTEIFKEHFSSYTFQAHSNSMPGREGVOLHPPLOOLSGFNGNSTHSDE----	766
Df	681	-----AMLGP-----SLSPQDPQEGGVYKHLHEKKYYRDD	712
Qy	767	-----LEYENOKLAEE--PEEDNLVLTFFEDLLCFAYQVAKGMFELEPK	808
Df	713	SGFSSQGVDTYVEMRPYSTNSDFSODLDKDGRLPLELDLHFSQVAQGNAFLASK	772
Qy	809	SCVHRDLAARNVLTGHGVKVICDFGLARDITLSDSSVYVRGNARLPVKWMAPELSLEGY	868
Df	773	NCHRDVAARNVLTNGHYAIGDFGLARDIMDSNYIKVGNARLPVKWMAPELSIEDCVY	832
Qy	869	TKSDWSYSGILLMEIFSGLVNPYPGIPVDANFYKLQSGFKMCPQPYATEGIIYFVQSC	928
Df	833	TQSDWSYSGILLMEIFSGLVNPYPGILVNSKFYKLVDGYQMAQPAFAPAKNIYSIQAC	892
Qy	929	WAFDSRKRPSPNLTSLFGCLQAEAEACIRTSHTLPKQAPQOQGLRAQSPQ	982
Df	893	WALEPTHRTPTQQICSLF--QEQAQERRDRDYTNLPSSS---RSGSGGSSES	941
RESULT 8			
KFMS_MOUSE	STANDARD:	PRT:	977 AA.
ID	KFMS_MOUSE		
AC	P09581; Q9DBH9;		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Macrophage colony stimulating factor I receptor precursor (CSF-1-R)		
DE	(EC 2.7.1.112) (fms proto-oncogene) (c-fms).		
GN	CSF1R OR CSFMR OR FMS.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	BEDLINE=88217329; PubMed=2966922;		
RX	Rochwell V.M.; Bourscheidt R.		

RX MEDLINE=94215924; PubMed=7512939;
RA Kubota T., Hikono H., Sasaki E., Sakurai M.;
RT "Sequence of a bovine c-kit proto-oncogene cDNA.";
RL Gene 141:305-306(1994).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: D16680; BAA04084.1; -
CC HSP: P11362; LFGE.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003598; Ig_c2.
CC InterPro: IPR003600; Ig_like.
CC InterPro: IPR001824; RTKinaseII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; Ig; 1.
CC Pfam: PF00069; pkinase; 1.
CC Pfam: PD000001; Euk_pkinase; 2.
CC SMART: SM00410; IG_like; 2.
CC SMART: SM00408; IGC2; 1.
CC SMART: SM00219; Tyrc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene: Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
CC Immunoglobulin domain.
FT SIGNAL 1 22
FT CHAIN 23 977
FT CHAIN 23 977
FT DOMAIN 23 521
FT TRANSMEM 522 544
FT DOMAIN 545 977
FT DOMAIN 590 938
FT NP_BIND 596 604
FT BINDING 624 624
FT ACT_SITE 793 793
FT MOD_RES 824 824
FT CARBOHYD 94 94
FT CARBOHYD 130 130
FT CARBOHYD 145 145
FT CARBOHYD 284 284
FT CARBOHYD 294 294
FT CARBOHYD 301 301
FT CARBOHYD 321 321
FT CARBOHYD 353 353
FT CARBOHYD 368 368
FT CARBOHYD 401 401
FT CARBOHYD 464 464
FT CARBOHYD 487 487
FT SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;

Query Match 32.1%; Score 1215; DB 1; Length 977;
Best Local Similarity 23.8%; pred. No. 8.1e-74;
Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;

QY 134 VTE---TOAGEYLLHIQSERANYTVLFTVNVRTQ-LVYLRPRYFRKMNQDALLCI-- 186
DB 83 ITEKAEATNTGNYTC---TNKGGLSSSIYVVRDPEKFLDLDPLYGKEENDTLVRCP 139
QY 187 -----SEGVEP-TVEWVLCSSHRESCKEKGPAVVRKEEVLHFGTDIRC-- 232
DB 140 DPEVTWYSLTGCSEKPLPKDLTEV-----ADPKAGITIRNVKREYHRL-----CLH 185
QY 233 CARNALGRE-CTKLTFTDLNQAPOS-----TLPQLFKVGEPLWIRCKAIHVNHGFLT 285
DB 186 CSANQKGSMLSKKFTILKVRRAIKAVPVSVSKTSYLLREGEFEFAVTLCKIDVSSVDSM 245
QY 286 WELEDKALEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPSSALVTI 345
DB 246 WIKENSQOTKATQKKNKSHWOGDFSYLRQERLTITISSARVNDSGVMCYANNTFGSANVTTT 305
QY 346 LE---KGFINA-TSSOEYELDPEKFCFVSRKAYPR-IRCTWIFSOASPPCQORGLD 400
DB 306 LEYVDKGFINFPMNTTVFVNDGENVDLVVEYEAYPEKPHRQWYMNRT----- 355
QY 401 GYSISKFCDHK-----NKPGEYIFVAENDDAQFTKMTNLNIRKK 439
DB 356 --STDKWDITPKSENENIRYVNELHLRLKGTGGGYTHVSNDSVNSVTFNVVNTK 413
QY 440 POVLANASA--SOASCSSDGYPLPSWTWKCKSDKSPNCTEEI-PEGVNNKANRKFQOW 496
DB 414 PEILTHDLVNGMLQCVAAAGPEPTIDWYFCPTGEQRCSVPVGVQVQIONSSVSPFGKL 473
QY 497 VSSSTLNMSEAGKLLVKCCAYNSMGTS-----CETIFLNSGPPFPFI 539
DB 474 VVYSTIDSTFKHGTVECRAYNDVGKSSASFNAFKGNSKEQIHAHTLFT-----PLL 527
QY 540 QDNISFATVGLCLPFIIVLVILCHYKKOFRYESQLOMI-QVTGPLDNEFYVDFRDY 598
DB 528 ---IGFVIAAGLMCIFVMIL-----TYKLLQKPYEVOVKVVEEING---NNVVIDPTOL 577
QY 599 EYDLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOQVAVKMLKEKADSCKEA 658
DB 578 PYDHKWEFPNRNLSFGKTLGAGFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREA 637
QY 659 LMSCLKMTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRKSRKFKHRTWTEI 718
DB 638 LMSCLKVLSVLGNHNVNLGACTIGTPTLVITEYCCYGDLLNLYLRKSRKFKHRTWTEI 690
QY 719 FKEHNFSSYTFQAHNSMMPGSRVQLPPLDQLSGFNGSIHSEDEI-----EYENQ 772
DB 691 -----ICSKQBDHAEVALYK-----NLLHSSKSSCNDSTNEYMDM 725
QY 773 K-----RLAEEEEEDLN-----VLTFEDLLCFAYQVAKGMEFLE 806
DB 726 KPGVSYVVPYKADKRRSARISYIERDVTYPAIMEDDELALDLEDLLSFYQVAKGMAFLA 785
QY 807 FKSCVHRDLAARNVLTGKVKVTKCDFGLARDILSDSYVVRGNARLPVKWMAPESLFEG 866
DB 786 SKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSYVYVGNARLPVKWMAPESEFNC 845
QY 867 IYTKSDVWSYGILLWELFSLGVNPPYPCIVDANFYKLIQSGFKMEQPFYATGTYFYMQ 926
DB 846 VYTFESDWSYGFELWELFSLGSSPYPGMPYDSKFYKMKIEGFRMLSPHAPAEYDINK 905
QY 927 SCWAFDSKRKPSFPNLTSLGCLQALAE 953
DB 906 TCWDADPLKRTFKQIVOLIEKQISES 932

RESULT 10
KFMS_RAT
ID KFMS_RAT STANDARD; PRT; 978 AA.
AC Q00495;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)

(EC 2.7.1.112) (fms proto-oncogene) (c-fms).
DE CSF1R OR CSFMR OR FMS.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RN SEQUENCE FROM N.A.
RC STRAIN=WiStar; TISSUE=Skeletal muscle;
RX MEDLINE=93001225; PubMed=1389227;
RA Borycki A.G., Guillier M., Leibovitch M.P., Leibovitch S.A.;
RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence
analysis and regulation during myogenesis.";
RL Growth Factors 6:209-218(1992).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X61479; CAA43706.1;
DR PIR: S16385; S16385.
DR HSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig_4.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00410; IG_like; 3.
DR SMART: SM00408; ICG2; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 978
FT MACROPHAGE COLONY STIMULATING FACTOR I
FT RECEPTOR.
FT
FT DOMAIN 20 511
FT TRANSMEM 512 536
FT DOMAIN 537 978
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 24 104
FT IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 107 197
FT IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 204 298
FT IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 299 397
FT IG-LIKE C2-TYPE DOMAIN 4.
FT DOMAIN 398 503
FT IG-LIKE C2-TYPE DOMAIN 5.
FT DOMAIN 580 914
FT PROTEIN KINASE.
FT NP_BIND 586 594
FT BINDING 614 614
FT ATP (BY SIMILARITY).
FT ACT_SITE 776 776
FT BY SIMILARITY.
FT DISULFID 42 84
FT DISULFID 127 177
FT DISULFID 224 278
FT DISULFID 417 483
FT MOD_RES 697
FT PHOSPHORYLATION (AUTO-) (IN VITRO)

FT MOD_RES 706 706 (BY SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (IN VIVO)
FT MOD_RES 807 807 (BY SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (IN VITRO)
FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;
Query Match 23.0%; Score 1210; DB 1; Length 978;
Best Local Similarity 33.5%; Pred. No. 1.8e-73;
Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;
QY 78 EAATVEAESGSLTQVQLATPGDLSCLWPKHSSLGCPHFHDLQNRGIVSMALINVTET 137
DB 36 ETVTLRCVSGSEWD-----GPISYWTLDPSGS-----TLTRNATFK 77
QY 138 QAGEY-LLHIQSERANVTVLFTVNRDTQLYV-----LRRPYFRKMNQDALL-CIS 187
DB 78 NTGTYRTELEDPMAGSTTI-----HLVYKDPAHSNWLLAOEVTVVEGOEAVLPC- 128
QY 188 EGVEPTVFWVLCSSHRESCKEKGPAVVRK-----EKKVLHLEFGTDIRCC 233
DB 129 --ITDPALK-----DSVSLMRGGRQVLRKTVYFFSAWRGFIIRKAKVL----DSNTVVC 177
QY 234 ARNALGRECTKL-FTIDLN---QAPOSTL-POFLXV-GEPLWIRCKAIHNVHGFGLTW 286
DB 178 KTVNGRESTSTGIWLKVNVRVHPPEPKLEPSKLVLRGEAAQIVCSATNAEVNVL 237
QY 287 ELEDKALE--EGSYFEMSTYSTNRTMIRILAFVSSVGRNDTGYTC--SSSKHPSQSALV 343
DB 238 KRGTGLEIFLNSDFQNYKKVRAL-----SLNAVDFQDAGIYSCVASNDVGRTRATM 291
QY 344 T--ILEKGFINATSQE-EYEIDPYEKFCFVRKAYPRIR-CTWIFSOAFCPCORLE 399
DB 292 NFQVVEASAYLNLTSQSLLOEVSVDLSLTVHADAVPSIQHYNNWYLGPFPE-DORKLE 350
QY 400 -----DGVSIKFCDH--KNKPEGYIFYAENDDAOFTKMTFLNIRKKPOVLNANASAO 450
DB 351 FITQRAIYRYTFKFLNRVKASEAGQYFLMAQNKAGWNLLTFELTRYPEYVSTVMDVN 410
QY 451 AS-----CSSDGYPLPSWTWKCKSDKSPNCTEIPGQVNVN---KKANKRVFGOWYSSSTL 502
DB 411 GSDVLFCDVSGVPOPSVTWMECRGHTDCDEAQAQALQVWNDTHPEVLSOKPKDKVLIQSQL 470
QY 503 NMSEAGKGLLVKCCAYNSMGTSCTEIFLNSPGPFPPFFIDNISFYATIGLC---LPFIIVL 559
DB 471 PIGTLKHNMTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDESILTPVVVACSVMSLLVLL 530
QY 560 IVLCHYKKKOFRYESQLOMTQ-VTGPLDNEYFYVDFRDYDLKWEPPREMLEFGKVLG 618
DB 531 LLLLLYKKQKPYQVRWKIIEYEG---NSYTFIDTQLPYNEKWEPPRNLOFGKTI 587
QY 619 SGAFGRVNNATAGISKTGVSIOQAVKMLKEKADCEKALMSKLMTHLGHHDHNVNL 678
DB 588 AGAFGVVEATAFGLGKEDAVLKVAVKMLKSTAHADAEKALMSKLMTHLGHHDHNVNL 647
QY 679 LGACTLGPVYLIFEYCCYGDLLNVLRSKRE-----KFHRTWTBI 718
DB 648 LGACTHGGPVLVITEYCCYGDLLNVLRSKRE-----KFHRTWTBI 707
QY 719 FKEHNFSSYPTFOAHNSNMPGSRVQLPPLDQLSGFNNGSIHSEDEIYEYENKRLAE 778
DB 708 RRDGSGS-----SOGVDYVEMRPVYSTSSDSFFKOD-----L 740
QY 779 EEDLNVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARD 838

```

Db 741 DREPSRPLELWDLHFFSSVQACMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800
Qy 839 ILSDSYVVRGNARLPVKWAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVD 898
Db 801 IMNDSYVVKGNARLPVKWAPESILYVTQSDVWSYGILLWEIFSLGLNYPGILVN 860
Qy 899 ANFYKLIQSGKMEQPFYATEGYFYVMQSWAFDSDKRSPFNLTSLF 946
Db 861 NKFYKLVKGYQMAQVPFAPKNIYSIMQSCWDLPTERRPTFQOICPLL 908

RESULT 11
KIT_CANFA
ID KIT_CANFA STANDARD; PRT; 975 AA.
AC O97799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99142897; PubMed=9989791;
RA Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
RT "Clustering of activating mutations in c-KIT's juxtamembrane coding
RT region in canine mast cell neoplasms.";
RL J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY..
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF044249; AAD02327.1;
DR HSSP; P11362; lFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
Transferase; Glycoprotein; Phosphorylation; ATP-Binding; Signal;

```

```

KW Immunoglobulin domain.
FT SIGNAL I 24
FT CHAIN 25 975
FT DOMAIN 25 519
FT TRANSMEM 520 542
FT DOMAIN 543 975
FT DOMAIN 588 936
FT NP_BIND 594 602
FT BINDING 622 622
FT ACT_SITE 791 791
FT MOD_RES 822 822
FT CARBOHYD 96 96
FT CARBOHYD 132 132
FT CARBOHYD 147 147
FT CARBOHYD 286 286
FT CARBOHYD 296 296
FT CARBOHYD 303 303
FT CARBOHYD 355 355
FT CARBOHYD 370 370
FT CARBOHYD 403 403
FT CARBOHYD 466 466
FT CARBOHYD 489 489
FT SEQUENCE 975 AA; 109335 MW; 8F570BDB9F05B1CB CRC64;

Query Match 22.9%; Score 1207; DB 1; Length 975;
Best Local Similarity 31.9%; Pred. No. 2.8e-73;
Matches 330; Conservative 165; Mismatches 375; Indels 166; Gaps 34;

Qy 31 LPYIKCVLISHENNGSSAGKPPSSYRMVGRSGPEDLOCPTRROSEGTVYEAATVVAESGSI 90
Db 11 LCVLLLLLLGLVGTGSS--QPS-----VSPGSPSLPSIHPAKSELIVSGD- 54
Qy 91 TLVOLATPGLDLSCLVWFKHSSLCGQCFHDLQNRGIYSMAILNVTTETQAGEYLLHIQSER 150
Db 55 ELRLSCTDPGEVK--WTFE--TLG-QLNENTHNEWITEKA-----EAGHTGNVCTNR 102
Qy 151 ANYTVLFTVNVRO--TQLYVLRPRYFRKMENQDALLCISEGVPEPTVEVWLSCSHRESCKE 209
Db 103 DGLSRSYVTVRDPKFLVDLPYLGKGNLTVRC-----PLTDEVTNYSYLRGC-- 153
Qy 210 EGPVAVRKEEKVLHLEFGTDIR-----C--CARNALGRE-CTKLFTIDLNAPOS- 256
Db 154 EGRPLPKDLTFVADPKAGITIRNVKRYHRILCLHCSADQKGTVLSKFTLVKRAIRAV 213
Qy 257 -----TLPLQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFEMSTYSTRMTI 311
Db 214 PVSVSKTSSLLKGEAFVSMCFIKDYSSFVDSMWIKENSQQTNAQTOSNSHHGDFNFE 273
Qy 312 RILLAFVSSVGRNDTGYTYTSSSKHPQSALVTILE----KGFINA-TSSQEEYEDPYEK 367
Db 274 ROEKLIISARVNDSGVFMVCYANNTFGSANVTTTLEVVVDKGFINIFPMWSTTIFYNDGON 333
Qy 368 FCFSVRKAYPR-IRCTWIFSQASF--PCEQGLEDDGYISIKFCDH-----KNKPEY 417
Db 334 VDLIVEYEAYPEHQOQMIYNNRTFTDKWEDYPKSDNSNRIYSELHLTRLKGEGTY 393
Qy 418 IFYAENDDAQFTKMTLNRKKQVLANASQA--SCSSDGYPLPSWTWKCSKDSKSPNC 475
Db 394 TFQVNSDVNSSVTFNVYVNTKPEILLTHESLTNGMLQCWVAGFPEPAVGWYFCPGAEOEC 453
Qy 476 TEEI-PEGVWNKKANRVFGOWVSSSTLNMSEAGLLVKCCAYNSMGTSCTETFLNSPG 534
Db 454 SVPIGPMQVQMONSSLSPSGLVQSSIDYSFAFKHNGTVECRAYNNVGRS--SAFEF--- 508
Qy 535 PFPIQDNIIFYATIGLCLPFIV-----VLVLICHYKKYKFRYESOLOMI-OVTGPLD 587
Db 509 -FAPKEQIHPTLFTPLLIGFVIAGMCMCIIVMILTYIKLPKPMYEVQKVVVEEING--- 564
Qy 588 NEYFVVDPRDYEDLKWEPFRENLEFGKVLGSGAFGRVNNATAYGISKTGVSQIVAVKML 647
Db 565 NNYVIDTQPLDQKWEFPNRRISFGTLGAGAFGKVVATAYGLIKSDAAMTAVKML 624
Qy 648 KEKADSCKEALMSELKMMNTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLRSK 707

```

```
Db 625 KPSAHLTEREALMSKLVLSYLNHNMNIYNLLGACVGGPTLVITEYCCYGDLLNLRK 684
Qy 708 REKFHRTWTEIFKEHNFSSYPYFQAHSSSSMPGSRVQLHPPLDQLSGFNGSIHSEDEI 767
Db 685 RDSFICSKQE---DHG-----EVALYK-----NLLHSKSS 712
Qy 768 -----EYENOK-----RLAEBEEDLN-----VLTFEDLLCPA 795
Db 713 CSDSTNEYMDMPGVSYVYVTPTRADKRARSARIGSYIERDVTYPAIMDEDLLELLSFS 772
Qy 796 YQVANGMEFLEPKSVHRDLARNVLVTHGKVVVKICDFGLARDILSDSYVVRGNARLPV 855
Db 773 YQVAKGMFLASKNCTHRDLARNILLTHGRITKICDFGLARDIKNDYVVKGNARLPV 832
Qy 856 KWMAPESLPEGYITTKSDVMSYGLLWETFLSGVNPYGPVDPANFYKLIOGFKMEQPF 915
Db 833 KWMAPESIFNCVITFESDVMWSYGIWLWELFSLGSSYPGMPVDSKRYKMKIEGFRMLSPE 892
Qy 916 YATEGIYFMQSCAFDSCRKPSFNLTSFLGCLQALAEAEACIRTSIHLPKQAAPQORG 975
Db 893 HAPAEYDIMKTCWDADPLKRTFTKQIVOLIEKQISDSIN-----HIYSNLAN----- 940
Qy 976 LRAQSPQROVKTHRR 991
Db 941 -CSPNPERPVVDHSVR 955

RESULT 12
KIT_HUMAN
ID KIT_HUMAN STANDARD; PRT; 976 AA.
AC P10721;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Term placenta, and Fetal brain;
RX MEDLINE=86111521; PubMed=2448137;
RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Munemitsu S.,
RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;
RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine
RT kinase for an unidentified ligand.";
RL EMBO J. 6:3341-3351(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93064697; PubMed=1279499;
RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;
RT "Organization and nucleotide sequence of the human KIT (mast/stem
RT cell growth factor receptor) proto-oncogene.";
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
RN [4]
RP VARIANT LYS-583.
RX MEDLINE=92291284; PubMed=1376329;
RA Fleischman R.A.;
RT "Human piebald trait resulting from a dominant negative mutant allele
RT of the c-kit membrane receptor gene.";
RL J. Clin. Invest. 89:1713-1717(1992).
RN [5]
RP VARIANT LEU-584.
RX MEDLINE=92133600; PubMed=1370874;
RA Spritz R.A., Giebel L.B., Holmes S.A.;
RT "Dominant negative and loss of function mutations of the c-kit
RT (mast/stem cell growth factor receptor) proto-oncogene in human
RT piebaldism.";
RL Am. J. Hum. Genet. 50:261-269(1992).
RN [6]
RP VARIANT ARG-664.
RX MEDLINE=92020918; PubMed=1717985;
RA Giebel L.B., Spritz R.A.;
RT "Mutation of the KIT (mast/stem cell growth factor receptor)
RT protooncogene in human piebaldism.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
RN [7]
RP VARIANT VAL-816.
RX MEDLINE=94013473; PubMed=7691885;
RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,
RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,
RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;
RT "Identification of mutations in the coding sequence of the proto-
RT oncogene c-kit in a human mast cell leukemia cell line causing
RT ligand-independent activation of c-kit product.";
RL J. Clin. Invest. 92:1736-1744(1993).
RN [8]
RP VARIANTS PIEBALDISM GLY-791 AND VAL-812.
RX MEDLINE=93322624; PubMed=7687267;
RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;
RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)
RT proto-oncogene in human piebaldism.";
RL J. Invest. Dermatol. 101:22-25(1993).
RN [9]
RP VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.
RX MEDLINE=96287384; PubMed=8680409;
RA Riva P., Milani N., Gandolfi P., Larizza L.;
RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large
RT Italian kindred with piebaldism.";
RL Hum. Mutat. 6:343-345(1995).
RN [10]
RP VARIANT GIST VAL-559 DEL.
RX MEDLINE=98361155; PubMed=9697690;
RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,
RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,
RA Kitamura Y.;
RT "Familial gastrointestinal stromal tumours with germline mutation of
RT the KIT gene.";
RL Nat. Genet. 19:323-324(1998).
RN [11]
RP FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
RP GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
RP BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
RP ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
RP 3-KINASE (PI3K).
RN [12]
RP CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
RP tyrosine phosphate.
RN [13]
RP SUBCELLULAR LOCATION: Type I membrane protein.
RN [14]
RP DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL
RP ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC
RP DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF
RP WHITE SKIN AND HAIR THAT LACK MELANOCYTES.
RN [15]
RP DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL
RP STROMAL TUMOR (GIST).
RN [16]
RP SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
RP PROTEIN KINASES.
RN [17]
RP SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
RN [18]
RP DATABASE: NAME=PROW; NOTE=CD guide CD117 entry;
RP WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm".
RN [19]
RP DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
RP WWW="http://www.infobiogen.fr/services/chronocancer/Genes/KITID127.html".
RN [20]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
```



```

Db 779 GMAFLASKNCIHRDLAARNILLTHGRITICDFGLARDIKNDNSYVYVKGARLPVKWAP 838
QY 861 ESLEFEGIVTKSDVWSYGLLWEISLGVNYPGIPVDANFYKLIOSGKMEOPFFATEG 920
Db 839 ESIFNCVTFEEDSVWSYGLLWEISLGVNYPGIPVDANFYKLIOSGKMEOPFFATEG 898
QY 921 IYFVMSQWAFDSRKRSPNLTSLFGLQALAE 953
Db 899 MYDKMKTCDWADPLKRPKQIVOLIEKOISE 931

RESULT 13
KIT_CHICK
ID KIT_CHICK STANDARD; PRT; 960 AA.
AC 008156;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN Kit.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Brain;
RX MEDLINE=93292995; PubMed=7685729;
RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,
RA Naito M., Sakurai M.;
RT "Cloning and expression of the chicken c-kit proto-oncogene.";
RL Gene 128:257-261(1993).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO
CC PRESENT IN THE BURSA OF FABRICIUS, HEART, KIDNEY, LUNG, SPLEEN
CC THYMUS AND OVARY.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL; D13225; BAA02506.1;
CC PIR; JN0677; JN0677.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR001824; RTKinaselII.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; Ig_3.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00408; Igc2; 1.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

```

```

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 960 FT MAST/STEM CELL GROWTH FACTOR RECEPTOR.
FT DOMAIN 25 505 FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 506 530 FT POTENTIAL.
FT DOMAIN 531 960 FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 575 913 FT PROTEIN KINASE.
FT NP_BIND 581 589 FT ATP (BY SIMILARITY).
FT BINDING 609 609 FT ATP (BY SIMILARITY).
FT ACT_SITE 777 777 FT BY SIMILARITY.
FT MOD_RES 808 808 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 76 76 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 135 135 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 149 149 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 269 269 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 286 286 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 306 306 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 318 318 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 338 338 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 343 343 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 356 356 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 453 453 FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 469 469 FT N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;

Query Match 22.4%; Score 1181; DB 1; Length 960;
Best Local Similarity 31.8%; Pred. NO. 1.5e-71;
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;

QY 45 GSSAGRPSSYRVGRGSPEDLOCTPRQSEGVYEAATVEAESIYLQVLATPGDLSG 104
Db 24 GSPVHESSLVNKGELRLKCN-----EEGPTVWNFQNSDPSAKTR 65
QY 105 L-----WVKHSLGCPHFQDLQNRGIVSMALNVNTEQAGEYLLHIOSEANVTVLFTV 159
Db 66 ISNEKEWHHTKNATIRDIGRYECKSG-----SIVN-----SFYV 99
QY 160 NVADTOLYVLRPYFRKMENQDALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKE 219
Db 100 FVKDPNVFLVDSLIYCKEDSDILLVCLTPDP-VLNFTRKCDGKPLKNTMIPNPK 158
QY 220 ----KVLHELFGTDIRCCAR-NALGRECTKLFTIDLNQAP-QSTLPOL-----FLKVG 266
Db 159 GILIKNVQSRFKGQYQCLAKHNGVEKISEHIF---LNVRPVHKALPVITLKSVELLKEG 215
QY 267 EPLWIRCKATHVNHGFLTWELDEKALEGSEYFEMSTYSTRNTRIRILLARFVSSVGRNDT 326
Db 216 EEFVTCIITDVSSVKASWISYKSAIVTSKRNIGDYGYERK-----LTNIRSVGVNDS 271
QY 327 GYITCSSSKHP--SQSALVTI--LEKGFINATSSQE-EYEIDPYEKFCFSYRFRKAYPR 381
Db 272 GEFTC-QAENPFCKTNATVTKALAKGFVRLFATMTNTIDINAGQNGNLTVEYEAYPK-- 328
QY 382 CTWIFSOAEPCCQRGLEDGYSISKFDHKNK-----PGEVIF 419
Db 329 -----PKEEVVMYMNNTLQNSDSHYVFKFTVGNNSYTSSELHLTRLKGTEGGIYTF 378
QY 420 YAENDAAQFTKMTLAIIRKKPOVLANASASQ--ASCSSDGYPLPSWTWKCKSDKSPNCTE 477
Db 379 FVNSDASSSVTFNVYVTKTPEIITLDMGLNDILQCVATGFPAPTIIWYFCPGTEQRCILD 438
QY 478 E----IPEGVWNKKANRKY--FGQWVSSSTLNMSBAGLLVCCAYNSMGTSCETIFLNS 532
Db 439 SPTISPMDVKVSYTNSVSPSFERILVESTYNAS-MFKSTGTICCEASNSGDK-SSVFENF 496
QY 533 PGPFPTIQDNISFYA-----TIGLCPLFIVLIVLICHYKQFRYESOLOMI-QVTG 584
Db 497 A-----IKEQIRTHTLFTPLLIAFGAAGLMCIIVMLVYIYLOKPKYEVQWKVVEEING 551
QY 585 PLDNEYFYVDPRDYEDLKWEPFRENLEFGKVLGSGAFGRVMNATAYGISKTGSIQVAV 644

```


Db 552 --NNVYIDPTQPDYDKWEPNRRLSFGKTLGAGFGKVVVEATAYGLFKSDAAMTAV 608
Qy 645 KMLKEKADCKEALMSKLMWTHLGHNDINVLNVLGACTLSQPVVLIIEYCCYGLLNV 704
Db 609 KMLKPSAHLTEALMSKLSYLGNIINVLNVLGACTLGPTLVITEYCCYGLLNV 668
Qy 705 RSKREKF---HRTWTE--IFKEHNFSSVPTFOA-HSNMSPGSRREVLPDQLDLSGFN 757
Db 669 RKRDSDFCPCHEEHAEEAAYENLLHQEPTADAVNEYMDKPGVSYAVPPKADKKRPVK 728
Qy 758 GNSIHEDEIEYENOKRLAEEDENLVTFEDLCLCFAYQVAKGMEFLFKSCVHRLDAA 817
Db 729 SGS-----YTDQDVTLSMLEDELALDVEDLLSFSYQVAKGMSFLASKNCLHRDLAA 780
Qy 818 RNVLVTHGVKVICDFGLARDILSDSSVYVGRNARLPVKWMAPESEFEGYTIKSDVWSY 877
Db 781 RNLLTHGRITKICDFGLARDIRNSYVYKGNARLPVKWMAPESEIFNCVYTFESDVWSY 840
Qy 878 GILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIYFVMQSWAFDSRKR 937
Db 841 GILLWELFSLGSPFGMPVDSKFKYMIKEGYRMFSPECSPPEMYDIMKSCWDADPLQRP 900
Qy 938 SFPNLTSLFCQLAE 952
Db 901 TFKQIVOLIEQLSD 915
RESULT 14
KIT_FELCA STANDARD; PRT; 978 AA.
AC Q28889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=95140426; PubMed=7530827;
RA Herbst R., Munemitsu S., Ullrich A.;
RT "Oncogenic activation of v-kit involves deletion of a putative
RT tyrosine-substrate interaction site.";
RL Oncogene 10:369-379(1995).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S76596; AAB33207.1; .
CC HSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR001824; RTKinaseI1.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_IIL; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL; 1 22
FT CHAIN; 23 978
FT DOMAIN; 23 521
FT TRANSMEM; 522 544
FT DOMAIN; 545 978
FT DOMAIN; 590 938
FT NP_BIND; 596 604
FT BINDING; 624 624
FT ACT_SITE; 793 793
FT MOD_RES; 824 824
FT CARBOHYD; 94 94
FT CARBOHYD; 145 145
FT CARBOHYD; 284 284
FT CARBOHYD; 294 294
FT CARBOHYD; 301 301
FT CARBOHYD; 321 321
FT CARBOHYD; 353 353
FT CARBOHYD; 368 368
FT CARBOHYD; 401 401
FT CARBOHYD; 464 464
FT CARBOHYD; 487 487
SQ SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;
Query Match 22.4%; Score 1178.5; DB 1; Length 978;
Best Local Similarity 31.8%; Pred. No. 2.3e-71;
Matches 313; Conservative 162; Mismatches 373; Indels 143; Gaps 30;
Qy 36 CVLISHENNGSSAGKPSRYMRVSGSPEDLOCTPRRQEGTVYEAAATVEAESGSITLQVQ 95
Db 12 CVLLLLLRVOTGSSQPSA-----SPGWSLPSIH-PATSELIYVSGDEIRLL 57
Qy 96 LATPGDLSCLVFKHSSLCGQPHFDLQNRGIVSMALNVVTQAGEYLLHITQSERANVTV 155
Db 58 CTDPGFVK--WTPE--TLG-QSSEITHNEWITEKA----EATNTGNYTC---TNGGGLSS 105
Qy 156 LFTVNVDR-TOLYVLRPYRKYKMNQDALLCI-----SEGVP----- 193
Db 106 SIYFVRDPARKFLVDLPYKEDHDLVRCPLTDPEVTNYSLRGCEGKPLPKDLTFTVD 165
Qy 194 -----TVEWVLCSSHR--ESCKEKGPAVVRKEVHLFCTDIRCCARNALGRECTKLF 246
Db 166 PKAGITRNKVRHYRCLHCSAD-----RGKSVLSKKFKILKYRAIR-----AVP 212
Qy 247 TIDLNOAQPSLTPLQFLVKVGEPLWIRCKAIHVNHGFGTLWELEDKALEEGSYFEMSTYST 306
Db 213 VVSYSKASH-----LLREGEFEFSVMCLIKDYSSSDVSDMSWIKENSPTNAQPSNSHQG 266
Qy 307 NRTMIRILLARVSSVGRNDTGYTCSSSKHPSQSALVILE---KGFINA-TSSOEVEI 362
Db 267 DFNVEVROBLTISRANVDNSGVFMFCYANNTFGSANVTTLLEVAKGFNIFPMMNTTIFV 326
Qy 363 DPYEKFCFSEVFKAYPR-IRCTWIFSQASFCEQRLGLEDGYSISKFCDHK----- 411
Db 327 NDENVNLDVIEYEAIPKPEHORWVIM-----NRTLTDKEDYPKSDNESNIRYVSELH 379

```
Oy 412 -----NKPGEYIFAENDDAOETKMTLNRKPKOVLANAS--ASOASCSDDGYPLPSW 463
Db 380 LTRLKNGEGYTFQVSNDSVNSVTLNYYVNTKEILTHESLVSGILOCLVAGPEPTV 439
Oy 464 TWKCKSDKSPNCTEBI-PEGVWNKANKRVFQWYSSSTLNMSKAGLLVKCCAYNSMG 522
Db 440 DWYFCGAEQRCVPVGLDQVQMSVSPSKLVVQSSIDYSAFKHNGTVECRASNNGV 499
Oy 523 TSCETIFLNSGPPFFIQDN-----ISFATICGLCLPFIIVLVILVILCHKVK 568
Db 500 KT--SAFFN-----FAFGKNSKQMPHPTLFTPLLLGFIAGM---MCIIVMILTYKYL 549
Oy 569 KOFREYSOLOMT-QVTGPLDNEFYVDFRDEYDLKWEFPRENLEFGKVLGSGAFGRVNN 627
Db 550 QKPMYEVQKVVVEING---NNYVIDPTLPYDHKWEFPNRNLSFGKTLGAGAFKGVVE 606
Oy 628 ATAYGISKTGVSIOAVKMLKEKADSCKEALMSKLMTHLGHHDNIVNLLGACTLSGP 687
Db 607 ATAYGLIKSDAAMTAVKMLKPSAHLTEREALMSKLVLSYLCNHMNVNLLGACTVGGP 666
Oy 688 VYLIEFYCCYDGLLNLKSKREKF----HRTWTEIFKEHNFSSYPTFOAHNSMPSGSE 743
Db 667 TLVITEFYCCYDGLLNLKSKREKF----HRTWTEIFKEHNFSSYPTFOAHNSMPSGSE 721
Oy 744 -VOLHPPDLDSGNGNSTHSEDETEYENQKRLAEEDLNVLTDFEDLLCFAYQVAKGM 802
Db 722 YMDKPGSVYVPTKADKRRSARIGSYTERDVTYPAIMEDDELALDELSFSYQVAKGM 781
Oy 803 EFLEKSCVHRDLAARNVLVTHGKVKYKICDFGLARDILSDSSYVVRGNARLPVKWMAPE 862
Db 782 AFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKDNSYVVKGNARLPVKWMAPE 841
Oy 863 LPEGITYTKSDVWSGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQFPYATEGIY 922
Db 842 IFNCVYTFSDVWSGIFLWELFSGSSYPGMPVDKSFYKMEKGFRLMSPEHAPENY 901
Oy 923 FVMQSCWAFDSKRPSFPNLTSLFLGCOLAEA 953
Db 902 DIMKTCWDADPLKRTFKQIVOLIEKQISDS 932

RESULT 15
PGDS_RAT STANDARD; PRT; 1088 AA.
AC P20786;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha platelet-derived growth factor receptor precursor (EC 2.7.1.112)
DE (PDGF-R-alpha).
DE PDGFRA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=90220609; PubMed=2157969;
RA Lee K.H., Bowen-Pope D.F., Reed R.R.;
RT "Isolation and characterization of the alpha platelet-derived growth
RT factor receptor from rat olfactory epithelium.";
RL Mol. Cell. Biol. 10:2237-2246(1990).
[2]
RP SEQUENCE OF 33-524 FROM N.A.
RX MEDLINE=93305723; PubMed=8318539;
RA Herren B., Weyer K.A., Rouge M., Loetscher P., Pech M.;
RT "Conservation in sequence and affinity of human and rodent PDGF
RT ligands and receptors.";
RL Biochim. Biophys. Acta 1173:294-302(1993).
CC -!- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
```

```
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M63837; AAA40743.1; ALT_INIT.
CC EMBL; Z14118; CAA78488.1; -.
CC PIR; A34710; PFTGA.
CC HSP; P11362; IFCK.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001824; RTKinaseII.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; ig. 4.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00409; IG; 3.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein;
CC Transferrase; Phosphorylation; ATP-binding; Immunoglobulin domain;
CC Signal.
CC SIGNAL. 1 23
CC CHAIN 24 1088
CC
CC DOMAIN 24 524 ALPHA PLATELET-DERIVED GROWTH FACTOR
CC TRANSMEM 525 548 RECEPTOR.
CC DOMAIN 549 1088 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 592 953 CYTOPLASMIC (POTENTIAL).
CC NP_BIND 596 606 PROTEIN KINASE
CC BINDING 626 626 ATP (BY SIMILARITY).
CC ACT_SITE 817 817 ATP (BY SIMILARITY).
CC MOD_RES 848 848 BY SIMILARITY.
CC CARBOHYD 41 41 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC CARBOHYD 75 75 O-LINKED (POTENTIAL).
CC CARBOHYD 102 102 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 178 178 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 467 467 N-LINKED (GLCNAC...) (POTENTIAL).
CC CONFLICT 150 150 L -> R (IN REF. 2).
CC CONFLICT 519 519 S -> T (IN REF. 2).
CC SEQUENCE 1088 AA; 122641 MW; 590C8BB0418801E7 CRC64;
Query Match 22.0%; Score 1157.5; DB 1; Length 1088;
Best Local Similarity 32.3%; Pred. No. 6.7e-70;
Matches 300; Conservative 156; Mismatches 317; Indels 155; Gaps 29;
Oy 122 QNRGIVSMALNVNTEOAGE-----YLLHIQSERANVTVLFTVNVDRTOQLYLRPP--- 172
Db 74 ENNSLFTVTVLEVNVNAAHTGWTCYNTQTQTESE-----IEGRHIYIV---PPDP 124
Oy 173 -YFRKMNQDALLCISEG---VP-----EPTVEWLCSHR-----ESCKEEG 211
Db 125 MAFVPLGMDTSLVIVEDDSAIIPCLTDPDTEVLHNGRLVPASVDSRQGFNGTFSVG 184
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:26:24 ; Search time 70.9642 Seconds

(without alignments)
2860.307 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLVLSV.....RGGLRAQSPQVQVNIHRERS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1268	24.1	976	13 Q9W755	Q9W755 brachydanio
2	1234.5	23.5	974	11 Q63702	Q63702 rattus ratt
3	1232.5	23.4	978	11 Q63116	Q63116 rattus norv
4	1220	23.2	954	13 Q91909	Q91909 xenopus lae
5	1215.5	23.1	974	13 Q98503	Q98503 danio dangi
6	1215.5	23.1	977	13 Q98501	Q98501 danio nigro
7	1214	23.1	977	13 Q918N6	Q918N6 brachydanio
8	1209.5	23.0	979	6 Q8WN23	Q8WN23 canis fami
9	1208	22.9	964	6 Q97744	Q97744 sus scrofa
10	1207	22.9	964	6 Q9TQ01	Q9TQ01 sus scrofa
11	1207	22.9	964	6 Q9TQ00	Q9TQ00 sus scrofa
12	1205	22.9	978	6 Q9XS93	Q9XS93 canis fami
13	1204.5	22.9	977	13 Q98S07	Q98S07 danio kerri
14	1201	22.8	948	6 Q9TWD7	Q9TWD7 trichosurus
15	1198.5	22.8	977	13 Q98S04	Q98S04 danio albol
16	1194	22.7	972	4 Q99662	Q99662 homo sapien

17	1188.5	22.6	975	13 P79750	P79750 fugu rubrip
18	1173	22.3	945	6 O77589	O77589 equus cabal
19	1170.5	22.2	724	6 Q9MYN0	Q9MYN0 bos taurus
20	1136	21.6	1059	13 Q9DE49	Q9DE49 brachydanio
21	1126.5	21.4	1087	13 Q9PUF6	Q9PUF6 gallus gall
22	1110	21.1	1097	11 Q8R406	Q8R406 rattus norv
23	1073	20.4	1019	13 Q8UVR8	Q8UVR8 fugu rubrip
24	1060	20.1	1048	13 P79749	P79749 fugu rubrip
25	1013.5	19.3	986	13 Q8UVR9	Q8UVR9 fugu rubrip
26	997.5	18.9	923	6 Q97745	Q97745 sus scrofa
27	980	18.6	1379	13 P79701	P79701 coturnix co
28	974	18.5	1327	13 Q8QHL3	Q8QHL3 gallus gall
29	961	18.3	1363	11 Q91ZT1	Q91ZT1 rattus norv
30	957	18.2	1301	13 Q91ZU9	Q91ZU9 brachydanio
31	946	18.0	1173	13 Q9PTL0	Q9PTL0 brachydanio
32	939	17.8	563	11 Q925F7	Q925F7 rattus norv
33	937.5	17.8	1345	11 Q8VCD0	Q8VCD0 mus musculu
34	901.5	17.1	323	11 Q9EQ22	Q9EQ22 rattus norv
35	877.5	16.7	323	11 Q9EQ24	Q9EQ24 rattus norv
36	833	15.8	345	13 Q9PVU7	Q9PVU7 lethenteron
37	820.5	15.6	350	13 Q91416	Q91416 xenopus lae
38	798	15.2	796	13 Q91287	Q91287 pleurodeles
39	793	15.1	766	4 Q96KM2	Q96KM2 pleurodeles
40	793	15.1	785	4 Q96KML	Q96KML homo sapien
41	793	15.1	819	4 Q96KMO	Q96KMO homo sapien
42	791	15.0	824	13 Q90749	Q90749 gallus gall
43	790.5	15.0	806	13 Q90Z00	Q90Z00 brachydanio
44	790.5	15.0	820	4 Q96KL9	Q96KL9 homo sapien
45	788.5	15.0	922	13 Q90413	Q90413 brachydanio

ALIGNMENTS

RESULT 1

Q9W755	PRELIMINARY;	PRT;	976 AA.
ID	Q9W755		
AC	Q9W755;		
DT	01-NOV-1999 (TRENBLrel. 12, Created)		
DT	01-NOV-1999 (TRENBLrel. 12, Last sequence update)		
DT	01-JUN-2002 (TRENBLrel. 21, Last annotation update)		
DE	Kit receptor tyrosine kinase.		
DE	Kit.		
GN	Brachydanio rerio (Zebrafish) (Zebra danio).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99396707; PubMed=10393121;		
RA	Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;		
RT	"Zebrafish sparse corresponds to an orthologue of c-kit and is		
RT	required for the morphogenesis of a subpopulation of melanocytes, but		
RT	is not essential for hematopoiesis or primordial germ cell		
RT	development."		
RL	Development 126:3425-3436(1999).		
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-		
CC	PROTEIN KINASES.		
DR	EMBL; AF153446; AAD41890.1; -.		
DR	HSSP; P11362; IFCG.		
DR	ZFIN; ZDB-GENE-980526-464; kit.		
DR	InterPro; IPR000719; Euk_pkinase.		
DR	InterPro; IPR003598; Ig_C2.		
DR	InterPro; IPR003600; Ig_Like.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR001824; RTKaseIII.		
DR	InterPro; IPR001245; Tyr_pkinase.		
DR	Pfam; PF00047; Ig_3.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PD000001; Euk_pkinase; 2.		
DR	SMART; SM00408; IGC2; 1.		

DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 976 AA; 109278 MW; A80AA01658C1A2DA CRC64;

Query Match 24.1%; Score 1268; DB 13; Length 976;
Best Local Similarity 36.6%; Pred. No. 1.5e-102;
Matches 318; Conservative 138; Mismatches 302; Indels 110; Gaps 29;

QY 177 MENQDALLCISEGPEPVEVNVVWVLCSSHRESCKEEGPAV--VRKEKVLHLEHFGDIRCCA 234
DB 138 MKNLNQKRC--DGQPLPN-----SLRYSASLETGVSVQKVRKEFGECVVCVGTLDAAVTV 189
QY 235 RNALGRECTKLTITDLNOAPOSTLP-----QLFLKVGGEPLWIRCKAIHVNHGFLTWE 287
DB 190 KS--GR-----YQLTVRLVPPDAPPTITLQQRVLLTQGEKLSLCSSTSNVNSDLAVKWK 242
QY 288 LE--DKALEEGSFEMSTYSTNFTMIRILLAFVSSVGRNDTGYTCSSSKHPSSOSA--- 341
DB 243 APNGVNPVSHQNSHL---LTEPIHVRTAILSLSSVTMDQAGNYSCEAINEKGTAKPV 298
QY 342 LVTILEKGFNATS--SQEYEDIDPEKEFCFVRKAYPRIC--TWIFSOAEPCEORGL 399
DB 299 VWNIIYKGFNITSVDNSTRRVRAGESLSLRVMVMAYPKPHFTFSWSYS----- 346
QY 400 DGYSISKCDH-----KKNPGEYIFVAENDDAQFTKMTFLNIRK 438
DB 347 -GVKLTNTDHTVTRTHGNSVTSKLVRLKVSSEGIYTFSCLNDRATIRQTFFVHVIS 405
QY 439 KPQVLANASA--SQASCSGDGYPPLPSWTW-----KKCSKSPNCTEIPPEGVWNKKAN 489
DB 406 KPQIVSYEGPIDGQVRCVAGYPTQIKWYCDLPHSRCSNLL-NATQE-EEDVVTVTMT 463
QY 490 RKVFGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLSPGFPF--FIQDNISFY 546
DB 464 NPPFGKGAVERSLNTKNNYATLE--CVASANGEIYVTLFSISENTPVHELFPTLLIGV 521
QY 547 ATIGLCLEFIVLVILVILCHYKQKFRYESOLOMIOVTPGLDNEFYVDFRQYEDLKWFE 606
DB 522 AAVI-----LVILILVITTKYMKQKPIQIOWKVIE--GIHGNNYVIDPTQLPYDQWEE 575
QY 607 PRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSCLKMM 666
DB 576 PRDKLRFKGLTSGAGFKVVEATAYGMSKADVTMTVAVKMLKPSAHATEKEALMSCLKVL 635
QY 667 THLGHNDNTVNLGACTLSGPVYLFEYCCYGDLLNLYLSKREKFRFTWTEIFEKHNFS 726
DB 636 SYLGNHINTVNLGACTVGPTLVITEYCCFGDLLNLRRLRRRYFYFT---TLGEDAYVR 692
QY 727 YPTFAHNSMMPGSGREVOLHPD-LDOLSGFNSTHSEDEIYEENOKRLAEEDLNV 785
DB 693 NVNMOSEPNDSRNG--YMTMKPVSGLILSENRRSLNKGD--SYSDSDAVSEILQEDGLT 748
QY 786 LTFEDLLCFAYQVAKMEFEKSVHRDLAARNVLVTHGKVVKICDFGLARDILSDSSY 845
DB 749 LQTEDELLSFYSQVAKGMDFLASKNCIHRDLAARNILLTQGRVAKICDFGLARDITDSDNY 808
QY 846 VVRGNARLPVKWMAPELSPGIVTYIKSDVWSYGIILLWEIFSLGVNPNYPGIPVDANFYKLI 905
DB 809 VVKGNARLPVKWMSPEIFECVYTFESDWSYGIILLWEIFSLGSSPYPGMPVDSKFKYKMI 868
QY 906 QSGFKMEQPPYATGEGYFVNMQSWADSRKRPSPNLTSLFGCOLAFA-EACARTSIHL 964
DB 869 KEGYRMESPEFSPSEMYDINHSCWDADPVKRPSPFSKIVEKIEQOISDSTRHIYLNFSRL 928
QY 965 PKOAAPOQRGLRAQSPORQVKTHRRS 992

Db 929 P--AAP-----GPRESSH---VHRLNS 947

RESULT 2
Q63702 PRELIMINARY; PRT; 974 AA.
ID Q63702
AC Q63702;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE C-kit receptor tyrosine kinase isoform.
OS Rattus rattus (Black rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN/FMAl;
RT Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;
RA "Two isoforms of rat c-kit receptor tyrosine kinase."
RL Nucleic Acids Res. 0:0-0(0).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES
DR EMBL; X62491; CAA44354.1; -;
DR HSSP; P11362; IFCK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; Icc2; 1.
DR SMART; SM00410; IG_Like; 3.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;

Query Match 23.5%; Score 1234.5; DB 11; Length 974;
Best Local Similarity 32.3%; Pred. No. 1.3e-99;
Matches 323; Conservative 176; Mismatches 340; Indels 161; Gaps 37;

QY 36 CVLLISHENNGSSAGKPPSYRMVRGSPEDLOCTPRRQSEGTVEAATVEVAESGSTITLQVQ 95
DB 12 CVLLVLLRGOTGTSQSPA-----SGPEPSPSIQPAQSELIVEAGD-TIRLT 57
QY 96 LATPGDLSCLVFVKHSLGCPHFQDQNRGIVSMALNV-TETQAGEYLLHIQSERANYT 154
DB 58 CTDPAFVK--WTEE-----ILDVRIENKQSEWIR--EKAETH 92
QY 155 VLFT-----VNRD--TQLYVLRPRYPFKMENQDALLCISEGVPEPTV-ENVLC 200
DB 93 GKYTCVSGSLRSIYFVRDPAVFLVGLPLFGKENDALVRC---PLTDPQVSNYSLI 149
QY 201 SSHRESC-----KEEGPAVVRKEEKVLHLEFGTDIRCCARNALGRECTKLTITDLNQ 252
DB 150 ECDGKSLPTDLKFPNPKAGITIKNVRAYHRLC---IRCAAQREGKWMRSKDFTLKYRA 206
QY 253 A----POSTLPQL--FLKVGEPILWIRCKAIHVNHGFLTW-----ELEDKALEE 295
DB 207 AIKAIPTVSVPTSHLLKEGDTFVICTIKDVSTSVDSMWIKLNPPQSKAQVKKRNSWHQ 266
QY 296 GSYFEMSTYSTNFTMIRILLAFVSSVGRNDTGYTCSSSKHPSSOSAIVT---ILEKGF 352

```

Db 267 GDF-----NYERQETLT-----ISSARVNDSGVFMCVANNFTFGSANVTTLTKVVEKGFIN 316
QY 353 --ATSSQEEYEIDPYEKFCFSVRKAYPR-IRCTWIFSQASFCEQRGLE-----DGYSS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKPEHOQWLYMNT--PTNRGEDYVKSQNSNI 373
QY 406 KFCDH-----KNKPEEYIFAENDDAOFTKMTNIRKKPOVLA--NASASQASCS 456
Db 374 RYVNELRLRLKTEGTYTFLVNSDSVASVTFDYYVNTKPEILYDRLMNGRLQCVAA 433
QY 457 GYPLPSWTWKCKSDKSPNCTEETPE-GVNNKANRKFQGVQWSSSTLNMSEAGKGLLVK 515
Db 434 GPEPTIDWYFCTGABQRCRTVPVPPVDVQIONASVSPFGKLVQSSSDSSVERHNGTVE 493
QY 516 CAYNSMGTSCTIFLNSPGFPF-----IQDN-----ISFYATIGLCLPFIVVLIV 561
Db 494 KASNAVGS--SAFFN-----FAFKGSKKEIQIOPHTLFTPLLTGTVVYTAGL----MGIIVM 543
QY 566 KYKQOPRYESQLOMI-QVTGPLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSGAFGR 624
Db 544 KYLQKPMYEVQKVVVEING---NNVYIDPTQLPYDHWKWEFPRNRLSFGTKLGAGAFGK 600
QY 625 VNNATAYGTSKGVSTQVAVKMKKADCEKALMSELKMTHLGHHDNIIVNLLGACTL 684
Db 601 VYEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLKVLNHNHNIVNLLGACTV 660
QY 685 SGPVYLIFECYCYGDLNLRKREKFTHTWEIFEHNFSSYPTTQAHNSNMPGSGREV 744
Db 661 GGPTLVITEYCYGDLNLRKRSDF--IFSKEQEQADAALYKNLLHKSKESSCDSSNEY 718
QY 745 QLHPPLDQLSGFN-----GNSHSEDETEYENOKRLAEEEDLNLVTFEDILLCA 795
Db 719 -----MDKKPGVSYVVPKTKRRSARIDSYTERDVTAPIMDEDLALDL---EDLLSFS 770
QY 796 YQVAKGMFELEPKSCVHRDLAARNVLVTHGKVKYKICDFGLARDILSDSSYVVRGNARLPV 855
Db 771 YQVAKGMFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRNDSNVYVKGARLPV 830
QY 856 KMAPESLFEGLTYITKSDVWSGILLWELFSLGVNYPGIPVDANFYKLIQSGFKMEQPF 915
Db 831 KMAPESLFCNVYTFESDWSVGIFLWELFSLGSPYGPMPVDSFKYMKKEGFRMLSPE 890
QY 916 YATEGIYFVMSQWAFDSKRPSFNLTSLFGCLQAAEE 955
Db 891 HAPAAVEYEMKTCWDADPLKRTFTKQVVLIEKQISDSK 930

RESULT 3
Q63116
ID Q63116 PRELIMINARY; PRT; 978 AA.
AC Q63116;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C-kit receptor tyrosine kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=92003944; PubMed=1912577;
RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,
RA Morii E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;
RT "Characterization of ws mutant allele of rats: A 12-base deletion in
RT tyrosine kinase domain of c-kit gene.";
RL Blood 78:1942-1946(1991).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; D12524; BAA02094.1; -.
DR HSP; P11362; IFGK.

```

```

DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003598; Ig_G2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF000047; Ig; 1.
DR Pfam: PF000089; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; IG.1ike; 3.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
KW
SQ SEQUENCE 978 AA; 109341 MW; 0958C33F19889051 CRC64;

```

```

Query Match 23.4%; Score 1232.5; DB 11; Length 978;
Best Local Similarity 32.2%; Pred. No 2e-99;
Matches 323; Conservative 176; Mismatches 340; Indels 165; Gaps 37;

```

```

QY 36 CVLISHENNGSSAGKPSYRMVRGSPEDLOCTPRRQSEGTVEAATVEAEGSITLQVQ 95
Db 12 CVLLVLLRGQTGTSQPSA-----SPGEPSPPSIQPAQSELIVEAGD-TIRLT 57
QY 96 LATPGDLSCLVFKHSLGCPHFQDLQNRGIVSMAILNV-TETQAGEYLLHIQSERANYT 154
Db 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIR--EKAETHT 92
QY 155 VLFT-----VNVRD-TOLYVLRPYPRKMNQDALLCISGVPPEPTV-EKVL 200
Db 93 GKTYCVSGSLRSIYFVDRDPAVFLVGLPLGKENDALVRC---PLTDPOVSNVSLI 149
QY 201 SSHRESC-----KEEGPAVVRKEEVLHFLGTDIRCCARNALGRECTKLTIDLNQ 252
Db 150 ECDGKSLPTDLKFPVNPKAGITIKNVKRAYHRLC---IRCAAQREGKWMRSKFTLAVRA 206
QY 253 A---PQSTLPOL--FLKVGEPWIRCKATHVNHGFLTW-----ELEDKALEE 295
Db 207 AIKAIPVSVVPETSHLLKEGDTFTVICTIKDVSTSVDSMMIKLNPQPOSKAQVKNRSHQ 266
QY 296 GSYFEMSTYSTNRMTIRILLAFVSSVGRNDTGYTSSSKHPSQSALVT---ILEKGFIN 352
Db 267 GDF---NYERQETLT-----ISSARVNDSGVFMCVANNFTFGSANVTTLTKVVEKGFIN 316
QY 353 --ATSSQEEYEIDPYEKFCFSVRKAYPR-IRCTWIFSQASFCEQRGLE---DGYSS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKPEHOQWLYMNT--PTNRGEDYVKSQNSNI 373
QY 406 KFCDH-----KNKPEEYIFAENDDAOFTKMTNIRKKPOVLA--NASASQASCS 456
Db 374 RYVNELRLRLKTEGTYTFLVNSDSVASVTFDYYVNTKPEILYDRLMNGRLQCVAA 433
QY 457 GYPLPSWTWKCKSDKSPNCTEETPE-GVNNKANRKFQGVQWSSSTLNMSEAGKGLLVK 515
Db 434 GPEPTIDWYFCTGABQRCRTVPVPPVDVQIONASVSPFGKLVQSSSDSSVERHNGTVE 493
QY 516 CAYNSMGTSCTIFLNSPGFPF-----IQDN-----ISFYATIGLCLPFIVVLIV 561
Db 494 KASNAVGS--SAFFN-----FAFKGSKKEIQIOPHTLFTPLLTGTVVYTAGL----MGIIVM 543
QY 562 LICHYKKQFYESQLOMI-QVTGPLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSG 620
Db 544 VLAYKYLQKPMYEVQKVVVEING---NNVYIDPTQLPYDHWKWEFPRNRLSFGTKLGAG 600
QY 621 AFRGVNATAYGISKTGVSQVAVKMKKADCEKALMSELKMTHLGHHDNIIVNLLG 680
Db 601 AFGKYVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLKVLNHNHNIVNLLG 660

```

```
QY 681 ACTLSPVYLIFECYCGDLLNLYRSKREKPHRTWTETFEKHNFSYPTFOAHNSMSPG 740
DB 661 ACTVGGTLVITEYCCYCGDLLNLYRSKREKPHRTWTETFEKHNFSYPTFOAHNSMSPG 718
QY 741 SREVOLHPPLDOLSGFN-----GNSIHSEDEIEYENQKRLAEEEDLNVLTFEDL 791
DB 719 SNEY-----MDMKPGVSYVPTKTKDRKSARSIDTSIERDVTTPAIMEDEDELALD--EDL 770
QY 792 LCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKVICKDFGLARDILSDSSYVVRGNA 851
DB 771 LSFQVQVAKGMAFLASKNCIHRDLAARNLTTHGRITKICDFGLARDIRNDSNYVVKGNA 830
QY 852 RLPVKWMAPESEFEGYITTKSDWMSYIGILLWEIFSLGVNPDYGPVDFANFYKLIQSGFKM 911
DB 831 RLPVKWMAPESEFNCVYTESDWSYIGIFLWELFSLGSSPYGMPYDSKFYKMKREGFRM 890
QY 912 EOPFVATEGIFVWQSCWAFDSKRPSPNLTSLGCLQALAEAE 955
DB 891 LSPHAPAAAYEYMKTCWDADPLKRPFTKOVVOLIEKQISDSSK 934

RESULT 4
QY91909 PRELIMINARY; PRT; 954 AA.
AC Q91909;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE C-kit-related kinase 1 (Akrk1) precursor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8355;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=9534496; PubMed=7619732;
RA Baker C.V., Sharpe C.R., Torpey N.P., Heasman J., Wylie C.C.;
RT "A Xenopus c-kit-related receptor tyrosine kinase expressed in
migrating stem cells of the lateral line system.";
RL Mech. Dev. 50:217-228(1995).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
DR EMBL; Z48770; CAA88688.1; --
DR HSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00410; IG_like; 2.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Signal; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT SIGNAL 1
FT SIGNAL 19
FT SIGNAL 19
SQ SEQUENCE 954 AA; 106859 MW; 9748845CBE0B537D CRC64;

Query Match 23.2%; Score 1220; DB 13; Length 954;
Best Local Similarity 32.0%; Pred. No. 2.4e-98;
Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;

QY 77 YEAAATVEAESGSIPLQVLAATPPGLSC-----LWVFKHSLSGCQPHFDLONRIVSMA 130
DB 15 YTGDAVPKINDEDRVTYVNGDKVSLKEDRAHLVTLATQKSLMKKPR-DLKSRLP----- 68
```

```
QY 131 ILNVTET-----QAGEYLLHIQSERANVTLVFTVNRDQTLVLRPRPKME 178
DB 69 -LNNSETDOFFVIIKADLRHIGRYICTNTQENTSV--SLFKDPAFLDIPFDIVTE 125
QY 179 NODAL-LCISEGVPPTVEMVLCSSHRESCBEGPAVVRKEEKVLHELFGTDIRC----- 232
DB 126 GADTVGMCEP---TDPMDIAI-----EKC--DGSPLPE-----NFTTDTIEAGITIK 169
QY 233 -----CARNALGR-ECTKLTIDLNQAPQSTLQPLFL-----KVGEPLWIRC 273
DB 170 TVQLAFDSCYVCGSNKGTVKKSSSTFIHVKPVPK-VPTVFLSKSRQLVKTGPEFVTC 228
QY 274 KAIHVNHGGLTW-ELEDKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCS 332
DB 229 AVLDVFSTYKAQWLVDKGVTKOANRSSNVFSYNTLKS-----DGVPSRSRFTCQ 282
QY 333 SSKHPSQ---SALVTILEKGFINATSSQE-EYEIDPYEKFCFSVRPKAYPR-IRCTWIFS 387
DB 283 AENAIQGVNATFTLDIVDGYVNLTVLENTTISVNGADNLVKVYIDAYPHDPDGVWTFY 342
QY 388 QASFPCEQRLGEGYSISKPCDHKN-----KPEYIFYAENDDAQFTKMF 432
DB 343 NETL----LNTSDHYVATK--DEGNRYVSELHLIRLKGTEKGYVTFYTTNSDDASVSE 396
QY 433 TLNIRKKPQVLANASASQAS--CSSDGYPLPSWTWKKCDKSPNCTEIEPBGVWKK--A 488
DB 397 NIQVKTREPEILIAERTSEGTLOCVATGFPVPAIOWEFCPGSEORCTDYPLSPVNEAFIQ 456
QY 489 NRKVFQGWSSSTLNMSEAGKLLVKCCAYNSMGTS-----CETIFLNSPGP 535
DB 457 ENSLGRVIVESTIDVNDLKKNTGVQCVASNEVASVFSFAIKELRTHLTFT----- 511
QY 536 FPFQDNISEYATIGLCPLFIVVLIVLILCHIKYKQFYRESOLOMI-QVTGLDNEYFYVD 594
DB 512 -PLL---IGFIAAGL---MCIAVAVLMVKYLOKPKYEQWKVVEEING---NNYVID 560
QY 595 FRYEYDLKWEFPRENLEFGKVLGSGAGFVGMNATATGISTGYSIQVAVKMLKEKADSC 654
DB 561 PTQLPYDNKWEFPDRRLCFGKILGAGAGFKVVEATAYGLLKEDSRLTVAVKMLKPSAHT 620
QY 655 EKAALMELKMTHTLGHHDNIVNLLGACTLGGPVLYLFEYCCYGDLLNLYRSKREKPHRT 714
DB 621 EREALMSELKVLSTYVTHHKNIVNLLGACTVGGPVLVITEYCCYGDLLNLYRSKREK 677
QY 715 WTEIFKEHNFESSYPTFOAHNSMSPGSRVOLHPPLDOLSGFNHSIHSEDEIEYENQK- 773
DB 678 -----ICPKFEDNS-----EAALYKNL-----LNRDMGCEGMSEYIDHKP 713
QY 774 -----RLAEEEDLNVLTFEDLLCFAYQVAKGMEFLFKSC 810
DB 714 AVSYVPTKTKDRRSGSGFDQDVSVSIPPEDDL-ALDTEDLINFSYQVAGNMFASKNK 772
QY 811 VHRDLAARNVLTGHKVKVICKDFGLARDILSDSSYVVRGNAFLPVKWMAPESLFGIYTI 870
DB 773 IHRDLAARNLTTHGRITKICDFGLARDIRNDSNVVVKGNARLPVKWMAPESEFHCVYTF 832
QY 871 KSDVMSYGIILLWEIFSLGVNPDYGPVDFANFYKLIQSGFKMEQPFYATGFIYFVMSQWA 930
DB 833 ESDVMSYGIILLWEIFSLGSSYPRIPIVDSKFYKMKIDGYRMMSPECAPLEMYEIMRSCWN 892
QY 931 FDSRRKRPSPNLTSLGCLQALAEAE 954
DB 893 SDPLKRPTFKQIVQNVQEQQLSDSK 916

RESULT 5
QY98SU3 PRELIMINARY; PRT; 974 AA.
AC Q98SU3;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
```


DE Receptor tyrosine kinase Fms.
GN FMS.
OS Danio dangila.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=127599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
diversification in Danio.";
RL Dev. Genes Evol. 211:319-328(2001).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
CC EMBL; AF324479; AAK15301.1; -.
DR HSSP; P11362; 1FGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferrase; Transmembrane.
FT VARIANT 720 720 A -> T.
SQ SEQUENCE 974 AA; 109902 MW; DE7FD64FD0D78B37 CRC64;

Query Match 23.1%; Score 1215.5; DB 13; Length 974;
Best Local Similarity 34.6%; Pred. No. 6.3e-98;
Matches 339; Conservative 147; Mismatches 342; Indels 153; Gaps 37;

QY 54 YRMVRGSPDLQCTPR-RQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWFKHSS 112
Db 13 FQVOGWSE-----PRLSSGAL--ADTDVLDGSGS----PLQLCEGDPVTEF----- 57

QY 113 LGCQPHDLQNRGI-----VSMALNVTTQAGEY-LLHTQSRANVTVLFTVN 160
Db 58 ----PRLAKHKRYISKEVSKSRTERVEKATVDFT---GTYKCIYINGNDSNLSSVHF 109

QY 161 VRDQ-LYV-----LRRPFRKMNODALLCISEGPEPTVEWLCSSHRECKEEGP 214
Db 110 VDRSVLFVSPSTSLR--YVRK-EGEDLLPCLLTDPOAT-DFIFRMDNGSAAPYGMNAT 165

QY 215 VRKEKVL-----HELFGTDIRCCARNALGRECTKFTLDLNAQPOSTLPOLFK----- 264
Db 166 FDRKGVILRNVHGFNDYICSAIRGAERKESKIFSINVIQRLFP-PYVYLKRNYYK 224

QY 265 -VGEPLWIRCKAIHNVHGFGLTWELEDKALEGVSFYFSTNTMTIRILLAFVSSVGR 323
Db 225 LVGERLQISCTNNFNFNVTWTHSSRLPKAE--ERSTWEGDRLATESILT-IPSVOL 281

QY 324 NDTGYTCSSKHPQSQALVT---ILEKGF-----NATSSQEEYEIDPYERKCFSV 372
Db 282 SDTGNTCTGONEAGANSSTQLLVVDPEYIRLSKLSKLTNRDLSEVSGDDVDLGV 341

QY 373 RFKAY-PRIRTW----IFSQASFPCEQR--GLEDDYSISKFDHKN--KPGEYIFYAEND 424

Db 342 WICAYPLLIHKWETPTSHNASLP--ENRFYNHNDRIEALLFLKRLNFEIEIGQYTLNVKNR 400
QY 425 DAOFTKMTNIRKKQOVLAN-ASASQASGSDGYPPLSPMTWKKCSKSPNCTEE----- 478
Db 401 VKSASITFDIKMYTKPVARVKWENWTTLSCRSYGYPAPSIILWYQCTGIRTTCPENTTDLQ 460
QY 479 -----IPEGVNNKANRVFGQWSSSTLNMSSEAKGLLVKCCAYNSMTSCETIFLN 531
Db 461 PIQTHGGVPEG-----ESFGVGVESVLTG-PNRRMTVVCVAFNLVGGQSDTFESMD 511
QY 532 SPGPFPIQDNISFYATIGLCLPIV--VLIVLICHYKKQFRYESQLQMTQVTPLDNE 589
Db 512 -----VSQIITSACMGSTVAMVILGGLLIFMIYIKQKPRYEIRKIIEATN--GNN 562
QY 590 YFYVDFRDYEDLKWEPPRENLEFGVLGSGAFGRVNNATAYGISKTGVSIOQAVKMLKE 649
Db 563 YTFIDPTQLPYNEKWEPPRDKLKLGTGLGAGFGKVVEATAYGLGKEDNITRVAVKMLKA 622
QY 650 KADSCKEALMSLKMTHLGHHDNIVNLLGACTLSGPVYLIFYCCYGDLLNLYRSKRE 709
Db 623 SAHPDEREALMSLKLHLGQHKNIIVNLLGACTHGGPVLVITEYCCCHGDLNLFRSKAE 682
QY 710 KFHRTWTEI--FKE-----HNFSSYPTFOAHSSNMPGS-----REVQLHPPLDOLS 754
Db 683 NFLNFVMTIPDFPEPMTDYKNISTERMF--VRSDSGISSACSDHYLDMRPVTSRPTLSSS 741
QY 755 GFNGNSTHSEDEIYEYENQKRLAEEDLNVLTFEDLLCFAYQVAKGMFLEFKSVHRD 814
Db 742 -----ECQEDSWPLDMDLRLFSQVAAQGLDFLAAKNCHRD 778
QY 815 LAARNVLVTHGVKVKICDFGLARDILSDSSIVYVGNARLPVKWMAPELSEFEGYTIKSDV 874
Db 779 VAAARNVLLTNSRVAKICDFGLARDIMDNSNVYVGNARLPVKWMAPESEFCVYTVQSDV 838
QY 875 WSGVILLWEIFSLGVNPPGIPVDANFYKLIQSGFKMEQPFYATEGYFYVQSCWAFDSR 934
Db 839 WSYGIMLWEIFSLGSPYPNILDVSKFYKMKICGYQMSRPFDFAPPETMTMKMCWNLDAA 898
QY 935 KRPSFPNLTSFLGCOLAEAE 955
Db 899 ERTFSKISQMIQMLGETPE 919

RESULT 6
Q98SUL ID Q98SUL PRELIMINARY; PRT; 977 AA.
AC Q98SUL;
DT 01-JUN-2001 (TrEMBLrel. 17, Created).
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update).
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update).
DE Receptor tyrosine kinase Fms.
OS Danio nigrofasciatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=144739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
diversification in Danio.";
RL Dev. Genes Evol. 211:319-328(2001).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
CC EMBL; AF324481; AAK15303.1; -.
DR HSSP; P11362; 1FGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig-like.

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig_4.
DR Pfam: PF00069; kinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 4.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; IG-like; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane.
FT VARIANT 60 V -> L.
FT VARIANT 103 103 A -> S.
FT VARIANT 742 742 P -> S.
FT VARIANT 899 899 E -> D.
FT SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;
Query Match 23.1%; Score 1215.5; DB 13; Length 977;
Best Local Similarity 33.5%; Pred. No. 6.3e-98;
Matches 334; Conservative 160; Mismatches 356; Indels 147; Gaps 35;
QY 54 YRMVRSPELQCTPR-ROSEGVYEAATVEAESITLQVQLATPGDLISLW---VFK 109
DB 13 FGQVQWSE-----PRILNSGAL--AGTDVILDGSG--PLQLVCEGGDGPVTFPRVAK 62
QY 110 HSLGQCPHFQDLQNRGIVSMAILNVTETQAGEY-LLHIQSERANYTVLFTVNVNRTQ-LY 167
DB 63 H-----KRYISKVGVGRFVRKATVDFGTQYKCVINGSNLAASSVHFVRDYSVLF 117
QY 168 V-----LRPYFKMENODALLCISEGVPTVEWVLCSSHRESCKEKGPAVVRKKEKVL 222
DB 118 VSPSSLR--YVRK-EGEDLLLCLLTDPAT--DFTFMDNGSNAAPGMNATFDPKGV 173
QY 223 ----HELFGDIRCCANALRECTLFTIDLNAQPOSTLPQLFLK-----VGEPLWI 271
DB 174 IRNVHGFNADYICARIGAEKVSKIFSNVIOQLRFP--PYVVKRNEYVKLVGERLQI 232
QY 272 RKAHVNHGFGTLWELEKALGEGSFEMSTYSTNTMIRILLAFVSSVGRNDGTGYTC 331
DB 233 SCTTNPNFYINVTWTHSSKRLPKPE--EKSTMBGDLAIESILT--ISSVQLSDTGNITC 289
QY 332 SSSKHPQSALVT---TLEKGFNAT---SSOEY-----EIDPYEKFCFSVRFKAYPRI 380
DB 290 TGONEAGANSSTQLLVVDEPYIRLSKLSKLTGRLGSLIEVSAGEDVDLGLVLEAYPPL 349
QY 381 RC-TW---IFSQASFPCEQR--GLEDGYSISKFDHKN--KPGEYIFYAENDDAQTFKMF 432
DB 350 TSHTWETPTSHNASLP--ENRFYNHNDRYEALLFLKRLNFEIEIGQYTLNVKNSMKSASITF 408
QY 433 TLNTRKKPQVLAN-ASASQACSSDGGPLPSWTWKKSKDSKSPNCTEEI---PEGVNNKK 487
DB 409 DIKMYTRPVARKWENVTTLSCRSYGPAPPSILWYQCTGIRTTCPENTDLDQIOTOTVE 468
QY 488 ANRVFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCETIFLSPGPFPIQDNISFYA 547
DB 469 FQKESFGAVGVESVLTG--PNRRMTVVCVAFNLVGGQSDTFSMD-----VSQIFTS 520
QY 548 TIGLCLPFIV--VLIVLICHYKKQFYSQLOMIQVTPGLDNEYFYVDFRDEYDLKWE 605
DB 521 MCGSTVAMVLLGLLFIWYKKQPKRYEIRWKIEATN--GNNYTFDPTQLPYNEKWE 578
QY 606 FPRNLEFGKVLGSGAGFRVNNATAYGLSKTGVSVIOAVKMLKADCKEALMSLKM 665
DB 579 FPRDKLGLKTLGAGAFKGVVVEATAYGLGKEDNITRVAVKMLKADDEREALMSLKI 638
QY 666 MTHLGHHDNIVNLLGACTLSFPVLIFFEYCCYGDLLNLYRSKREK----- 711

DB 639 LSHLGQHKVIVNLLGACTHGGPVLVITECCRGDLNLFASKAENFLNFVMTIPNPEPV 698
QY 712 -----HRTW-----TEIFKEHNFSSYPTFQAHSSSMSPGSRREVQLHPDLQLSGF 756
DB 699 TDYKNVSTERMFVRSDSGISSTCSHYLDMPVTSRPTNSALDPSSDCQ----- 747
QY 757 NGNSIHSEDEIEYENOKRLAEEDDLNVLTFFDLICFAQYAKGMFELEFKSCVHRDLA 816
DB 748 -----EDSWPLDMDDLLRFSSQAQGLDLAAKNCIHRDVA 783
QY 817 ARNVLVTHGKVKYKICDFGLARDILSDSSVYVGNARLPVKWMAPESEFEGYITIKSDVWS 876
DB 784 ARNVLLTNSRVAKICDFGLARDIMNDSNVYVGNARLPVKWMAPESEFECVYTVQSDVWS 843
QY 877 YGILLWEIFSLGVNYPGPVDPANFYKLIQSGFKMEQPFYATEGIVYVMSQWAFDGRKR 936
DB 844 YGIMLWEIFSLGKSPYNILVDSFKYMKIKGQYQMSRPFAPPENYIMTKMKNLEAAER 903
QY 937 PSPNLTSLGCOLAE-AEACIRTSTIHLPKQAAPQ 972
DB 904 PTFKISQMIQRLGETSEQDTQYKNIPISEAEVEQ 940
RESULT 7
Q918N6 PRELIMINARY; PRT; 977 AA.
AC Q918N6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fms.
GN CSF1R OR FMS.
OS Brachydanio rerio (zebrafish) (Zebra dario).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB;
RX MEDLINE=20333166; PubMed=10862741;
RA Parichy D.M., Ransom D.G., Paw B., Zon L.I., Johnson S.L.;
RT "An orthologue of the kit-related gene fms is required for development
of neural crest-derived xanthophores and a subpopulation of adult
melanocytes in the zebrafish, Danio rerio."
RL Development 127:3031-3044(2000)
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
DR EMBL; AF240639; AAF76872.1; -.
DR HSP; P11362; IFGK
DR ZFIN; ZDB-GENE-001205-1; csf1r.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_4.
DR Pfam; PF00069; kinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transmembrane; Tyrosine-protein kinase.
FT VARIANT 238 238 N -> K.
FT VARIANT 615 615 V -> M.
FT SEQUENCE 977 AA; 110187 MW; C91A2F339E746A58 CRC64;

Q8WN23	PRELIMINARY;	PRT;	979 AA.
AC	Q8WN23;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	C-KIT.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxID=9615;		
RN	[1]		
RP	SEQUENCE FROM N A.		
RA	Zemke D., Yuzbasian-Gurkan V.;		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF448148; AAL40833.1;		
DR	InterPro; IPR000719; Euk_pkinase.		
DR	InterPro; IPR003599; Ig.		
DR	InterPro; IPR003600; Ig_like.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR001824; RTK_kinaseIII.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR001245; Tyr_pkinase.		
DR	Pfam; PF000047; Ig_2.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PD000001; Euk_pkinase; 2.		
DR	SMART; SM00409; IG; 4.		
DR	SMART; SM00410; IG_like; 2.		
DR	SMART; SM00220; S_TKC; 1.		
DR	SMART; SM00219; TYKIC; 1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.		
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.		
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; UNKNOWN_1.		
DR	SEQUENCE 979 AA; 109753 MW; 46C30D5DEB8E33D3 CRC64;		
QY	Query Match	23.0%;	Score 1209.5; DB 6; Length 979;
Db	Best Local Similarity	31.3%;	Pred. No. 2.1e-97;
Db	Matches 332; Conservative 165; Mismatches 369; Indels 195; Gaps		
QY	14 LNVLSVMILETVNQDLPVVKVLISHENNGSSAGKPFSSYRMVGRSPEDLOCTPRQSE 73		
Db	11 LCVLILLGLLVGTGSSQPSV-----SPGEPSSL-----		
QY	74 GTVVEAATVVAESGSIITLQVLATPGDLSCLVAFKHSSIGCPHFDLQNRGIVSMAILN 133		
Db	39 PSIHPAKSELIVSGD-ELRLCTDGFVK--WTFE--TLG-QLNENTHNEWITEKA--- 89		
QY	134 VTETQAGEYLLHTOSERANTVTLFTVNVRD-TOLYVLRPFYRKMNQALLCISGVPE 192		
Db	90 ----EAGHTGNYTCTNRDGLSRSIYFVVRDPAKFLFVLDPPLYCKEGNDLIVRC----- 138		
QY	193 PTVEWVLCSSHRESCEEGPAVVRKEEVLHFLGTDIR-----C-CARNALGRE 241		
Db	139 PLTDPPEVTNYSLRGC--EGKPLPKDLTFLVADPRAGITIRNVKREYHRLCLHCSADQKGT 196		
QY	242 -CTKLITDLNQAPOS-----TLPOLFLKVGPELIRCKA IHNVHGFGLTWELEDKALE 294		
Db	197 VLSKKFTLKVRAAIRAVPVVSVSKTSSLCKEGEAFVMCFIKDPFVDSVMWIKENSQOT 256		
QY	295 EGSYFEMSTYSTNRTMIRILLAFVSSVRGNDTCYGTCCSSKHPSQSALVTILE---KGFI 351		
Db	257 NAQTOSNWHGGDFNERQEKLIIISARVNDSGVFVICYANNFTGFSANVTTLLEVVDKGI 316		
QY	352 NA-TSSOEEYEIDPYEKFCFSVRFKAYPR-IRCTWIFSOASF--PCEORGLLEDGYISIKF 407		
Db	317 NIFPMWSTTIFVNDGENVDLIVEYEAYPKPEHQOWIYMNRTFTDKWEDYPKSDNESNIY 376		
QY	408 CDH-----KNKPGEYIFVAENDDAQFTKMTFLNTRKKPKQVLIANASA SOA--SCSSDGY 458		
Db	377 VSELHLTRLKNGEGGYTFQVSSNDSDYNSSTFNVYNTKTPPEILTHESLNGMLOCVVAGF 436		

```
OY 459 PLPSMTWKCKSDKSPNCTEEI--PEGVWKNKANKRVFGOWVSSSTLNMSAGKGLLVKCA 517
Db 437 PEPADVWTFCEPAGBQRCVPIGPMQVOMQNSLSGSKLVVOSSIDYSAFKNGITVECR 496
OY 518 YNSMGTSCETIFLNSPGPPFTIODN-----ISFYATIGLCPLFIYVLVILI 563
Db 497 YNNVGRS--SAFFN---FAFKGNSKEQIHPTLTFTPLLIGFVIAAGM---MCIIVMIL 546
OY 564 CHKYYKQFRYESQLQMI-QVTGPLDNEYFYVDFRDYEDLKWEEFPRENLEFGKVLGSGAF 622
Db 547 TYKYLQKPMYEQWKVEEING---NNYVIIDPTQLPYDHKWEEFPNRLSFGKTLGAGAF 603
OY 623 GRVMAATAYGISKTVGSIOVAVKMLKEKADSCKEALMSKELKMTHTLGHNDNIVNLGAC 682
Db 604 GKVEATAYGLIKSDAMTVAVKMLKPSAHLTEREALMSKELKSVLSYLGNNHMINVLGAC 663
OY 683 TLSGPLYLIFCYCCYGLDLNLYLRSKREKPHRTWTETFEKHNFFSSYPTFOAHSSNMPGSR 742
Db 664 TVGGPTLVITEYCCYGLDLNLYLRSKREKPHRTWTETFEKHNFFSSYPTFOAHSSNMPGSR 701
OY 743 EVOLHPLDOLSGFNSTHSEDEI-----EYENOK-----RLAEEB 779
Db 702 EVALYK-----NLLHSESSCSDSTNEYMDKMGVSYVYVPTKADKRRSARIGSYI 751
OY 780 EEDLN-----VLTFFEDLLCFAYOVAKWMEFLFKSCVHRDLAARNVLYTHGKVKKI 830
Db 752 ERDVTPAIMDEDLADLELLSFSYOVAKWMAFLASKNCIHRDLAARNVLYTHGKVKKI 811
OY 831 CDFGLARDILSSSYVVRGNARLPVKWMAPELSIFEGYITKSDVMSYGLLWEIFSLGVN 890
Db 812 CDFGLARDIKNSYVVRGNARLPVKWMAPELSIFEGYITKSDVMSYGLLWEIFSLGVN 871
OY 891 PYGPVPDANFYKLIOSGKMEOPFVATGTYFVMSQWAFDSKRSPNLTSLFGLCQL 950
Db 872 PYGPVPDANFYKLIOSGKMEOPFVATGTYFVMSQWAFDSKRSPNLTSLFGLCQL 931
OY 951 AEAEEACIRTSHTLPOAPOOGRGLRAOSPOROVKIHRR 991
Db 932 SDSTN-----HIYNSLAN-----CSPNRPVVDHVSVR 959

RESULT 9
O97744
ID O97744 PRELIMINARY; PRT; 964 AA.
AC O97744;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KIT1*0101.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
OR [1]
RN SEQUENCE FROM N.A.
RP STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RC MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural sciences, Uppsala, SWEDEN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AJ223228; CAA11196.1; --
```

```
DR HSP; p11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
KW Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964
SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

Query Match 22.9%; Score 1208; DB 6; Length 964;
Best Local Similarity 32.6%; Pred. No. 2.8e-97;
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

OY 38 LISHEHNGSAGKPSRYMVRGSPEDLOCTPRROSEGTVEATVEAEGSITLVOLA 97
Db 16 LLLRVOTGSS-QPSV-----SPEL-----SPSIHPKSELIVSAGD-EIRLCT 59
OY 98 TPGDLCLWVFKHSSLCOPHFDLNLRGIVSMALNTVETOAGVYLLHIOSERA----- 151
Db 60 DPG-SVKWTFE-TLG-----QSENTHAEWIV---EKAENNTG 93
OY 152 NYTVL-----FTVNVRTQ-LYVLRPRYFRKMNQDAL-----LCISE 188
Db 94 NYTCNEGGLSSYIVFVRDPEKLFVDPPLYGK-EDNALVRCPLTDPEVTNYSLTGCE 152
OY 189 GVPEP-TVVEWLCSSHRESCKEGPAVVRKEEVHLFEGTDIRC--CARNALGRE-CTK 244
Db 153 GKPLPKDLTEV-----ADPKAGITIRNVKREYHRL-----CLHCSANQGKSVLSK 198
OY 245 LFTIDLNQAPQS-----TLPQLFKVGEPLWIRCKAIHVNHGFGLTWELDKKALEGSY 298
Db 199 KFTLKVRAAIRAVPVAVKASVYLLRGEFEFVAVMCLIKDVSSSVDSWM--IRENSQTRAQV 257
OY 299 FEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSKSHPSQSALVTILE---KGFINA-T 354
Db 258 KRNSWHOGDFNLRQERLTITISSARVNSGVFMCYANNTFGSANVTTLTLEVVDKGFNIFP 317
OY 355 SSOEEYIDPYEPCFVSFRKAYPR-IRCTWIFSQ--ASPPCEORGLDGYISKFCDHK 411
Db 318 MNNTTFVNDGEDVDLIVEYEAYPEKHROWIMNRTATDKWEDYPKSENSERYVSEL 377
OY 412 N-----KPEYIFYAENDDAQFTKMTLNIRKKPOVLN--ASASQACSSDGYPLPS 462
Db 378 HLTRLKGTGTYFYLVSNADVNSVTFNVYVNTKPELLTHDRLMNGMLOCVAGFPPEPT 437
OY 463 WTNRKCKSDKSPNCTEEI-PEGVWKNKANKRVFGOWVSSSTLNMSAGKGLLVKCAYSNM 521
Db 438 IDWYFCPCTEORCSVPVGPVDVQIONSVSVPFGKLVITHSSIDYSAFKHNGVVECRAYNDV 497
OY 522 GTSCTETFLNSPGPPFTIODNISFYATIGLCPLPIV-----VLIVLICHYKKQFRYES 575
Db 498 GKS--SAFFN---FAFKEQIHPTLTFTPLLIGFVIAAGMCIIVMILTYLQKPMYEV 551
OY 576 QLQMI-QVTGPLDNEYFYVDFRDYEDLKWEEFPRENLEFGKVLGSGAFGRVMAATAYGIS 634
Db 552 QWKVVEEING---NNYVIIDPTQLPYDHKWEEFPNRLSFGKTLGAGAFKGVVEATAYGLT 608
OY 635 KTGVSIQVAVKMLKEKADSCKEALMSKELKMTHTLGHNDNIVNLGACTLSGPVYLIFEY 694
```

Db 609 KSDAAMTVAVKMLKPSAHLITERALMSLKVLSYLGHNHNIIVNLLGACTIGTPTLVITEY 668
 QY 695 CCYGDLLNLYRSKRREKFRHTWTETFEKHNFSYPTQAHSSNMPGSRVQLHPPLDQLS 754
 Db 669 CCYGDLLNLYRSKRREKFRHTWTETFEKHNFSYPTQAHSSNMPGSRVQLHPPLDQLS 728
 QY 755 GFNGNSIHSEDEIEYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRD 814
 Db 729 PTKADKRSARIGSYIERDVTPAIMEDELALDLEDLLSFSYQVAKGMAFLASKNCIHRD 788
 QY 815 LAARNVLVTHGKVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESEFEGYTIKSDV 874
 Db 789 LAARNILLTHGRITKICDFGLARDIKDSDNYVVKGNARLPVKWMAPESEFNCVYTFESDV 848
 QY 875 WSYGILLWEISLGVNPPGIPVDANFYKLIQSGFKMEQPEYATEGIYFVMOCSWAFDSR 934
 Db 849 WSYGIFLWELFSLGSSPYGMPVDSKFKYMKIEGFRMLSPHAPAEYDIMKTCWDADPL 908
 QY 935 KRSPFNLTSLGCOLAEA 953
 Db 909 KRPFQVQLIEKQISES 927
 RESULT 10
 Q9TQ01 PRELIMINARY; PRO; 964 AA.
 AC Q9TQ01;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Mast/stem cell growth factor receptor (Fragment).
 GN K1P1*0201.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RA Marklund S.;
 RL Thesis (1997), Department of Animal Breeding and Genetics,
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RX MEDLINE=98391767; Pubmed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
 RA Moller M., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 RL pig";
 RL Genome Res. 8:826-833(1998).
 CC -i- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -i- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC EMBL: AJ232329; CAall1197.1; ..
 DR HSP; P11362; IFGK
 DR InterPro; IPR00719; Euk_pkinase.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003600; Ig_like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001824; RTKinaseII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00410; IG_like; 1.
 DR SMART; SM00219; TyrcK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
 KW Transmembrane; Tyrosine-protein kinase.
 FT NON_TER 964
 SQ SEQUENCE 964 AA; 108315 MW; 996C3C46201358A8 CRC64;
 Query Match 22.9%; Score 1207; DB 6; Length 964;
 Best Local Similarity 32.6%; Pred. No. 3 5e-97;
 Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;
 QY 38 LISHENNNGSSAKPSSYRMVGRGSPEDLOCTPRROSEGTVYEAATVEVAESGISTLQVOLA 97
 Db 16 LLRLVQTGSS--QPSV-----SPEL-----SPSIHPAKSELIVSAGD-EIRLFT 59
 QY 98 TEGDLSCLWVFKHSSLCGPHEDLQNRGIVSMAILNVTQAGEYLLHIOSERA----- 151
 Db 60 DPG--SVKWTFE--TLG-----OLSENTHAEWIV-----EKAEAMNTG 93
 QY 152 NTVL-----FTVNVRDTO-LYVLRPRYPRKMNODAL-----LCISE 188
 Db 94 NYTCTNEGGLSSSIYFVRDPEKFLVLDPLYGK-EDNALVRCPLTDPVTNYSLTGCE 152
 QY 189 GVPEP-TVEWVLCSSHRESCKEEGPAVVRKEKVLHLEFGTDIRC--CARNALGRE-CTK 244
 Db 153 GKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLHCSANOGSKVLSK 198
 QY 245 LETIDLNQAPQS-----TLPOLFLKVGPELWIRCKAIHVNHGFGLTWELEDKALEGSY 298
 Db 199 KFTLVRAAIRAVPVAVSKASYLLREGEFAVMCLIKDVSSVSDSMW-IRENSQTKAQV 257
 QY 299 FEMSYSTNRTMIRILLAFVSSVGRNDTCYTCSSSKHPSQSALVILE---KGFINA-T 354
 Db 258 KNSWHQDGFNLOERLTISSARVNDSGVFMCIYANTFGSANVTTLLEVVDGFINIFP 317
 QY 355 SSOEYEIDPYKEKFCFSVRFKAYPR-IRCTWIFSQ--ASFPCQORGLDGYSTSKCDHK 411
 Db 318 MNNTTVFVNDGEDVDLIVEYEAYPEHQWIYMNRTATDKWEDYKPSENESIRIVSEL 377
 QY 412 N-----KPEGYIFYAENDDAQFTKMTFLNIRKKQVILAN--ASASQSCSSDGYPLPS 462
 Db 378 HLTRLKGTGGTYTFLVSNADVNSVTFNVVYNTKPEILTHDRMLMGLQCVAAGPPEPT 437
 QY 463 WTKKCDKSPNCTEEI-PEGVNNKANKRVFGQVSSSTLNKSEAGKGLLVKCCAYNSM 521
 Db 438 IDWYFCPTGTEQRCVPVGVVDVQIONSSVSPFGKLVHSSIDYSAPKHNGTVECRAYNDV 497
 QY 522 GTSCTEIFLNSPGPFPIQDNISFYATIGLCPLPFIV-----VLIVLCHKYKKQFRYES 575
 Db 498 GKS--SAFEN---FAKQEIHAHTLFTPLLGIVIAAGMMCIIVMLTYKYLQKPMYEV 551
 QY 576 QLOMI-QVTGPDLDNEYFYVDFRDEYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYGIS 634
 Db 552 QMKVVEEING---NNYVIDPTQLPYDHWKPEPRNRLSFGKTLGAGAFKVVEATAYGLI 608
 QY 635 KTVSYIQVAVKMLKEKADSCKEALMSELKMTWHLGHHDNIIVNLLGACTLSGPGVILIFEY 694
 Db 609 KSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNIIVNLLGACTIGTPTLVITEY 668
 QY 695 CCYGDLLNLYRSKRREKFRHTWTETFEKHNFSYPTQAHSSNMPGSRVQLHPPLDQLS 754
 Db 669 CCYGDLLNLYRSKRREKFRHTWTETFEKHNFSYPTQAHSSNMPGSRVQLHPPLDQLS 728
 QY 755 GFNGNSIHSEDEIEYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRD 814
 Db 729 PTKADKRSARIGSYIERDVTPAIMEDELALDLEDLLSFSYQVAKGMAFLASKNCIHRD 788
 QY 815 LAARNVLVTHGKVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESEFEGYTIKSDV 874
 Db 789 LAARNILLTHGRITKICDFGLARDIKDSDNYVVKGNARLPVKWMAPESEFNCVYTFESDV 848
 QY 875 WSYGILLWEISLGVNPPGIPVDANFYKLIQSGFKMEQPEYATEGIYFVMOCSWAFDSR 934
 Db 849 WSYGIFLWELFSLGSSPYGMPVDSKFKYMKIEGFRMLSPHAPAEYDIMKTCWDADPL 908

[illegible]

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RA London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,
RA Geissler E.N.;
RT "Spontaneous canine mast cell tumors express tandem duplications in
RT the proto-oncogene c-kit.";
RL Exp. Hematol. 0:0-0(1999);
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF090300; AAD28369.1; -.
DR HSP; P11362; IFCK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 2.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein Kinase.
SQ SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

Query Match 22.98; Score 1205; DB 6; Length 978;
Best local similarity 31.5%; Pred. No. 5.4e-97;
Matches 329; Conservative 167; Mismatches 387; Indels 160; Gaps 33;

QY 14 LNVVLSVILETVNODLPVVKLVLSHENNGSSAGPSSSYRMVGRGSPEDLOCTPRROSE 73
Db 11 LCVLLLLLVGVQSGSQPSV-----SPGPESL----- 38

QY 74 GTVEAATVEVAESGSIITVLQVATPGDLSCLVWFKHSSLCQPHFDLQNRGIVSMALN 133
Db 39 PSIHAKSELIVSGD-ELRLSCDTPGFVK--WTFE--TLG-QLNENTHWEIWEKA--- 89

QY 134 VTETQAGEYLLHIQSERANTVTLFTVNVRD-TQLYVLRPPYFRKMEQDALLCISEGVPE 192
Db 90 -----EAGHTGNTCTNRDGLSRISYFVROPAPKFLVDLPYLGKENDTLVRC----- 138

QY 193 PTVEWVLCSSHRESCKEGRPAVVRKEKVLHELFGTDIR-----C--CARNALGRE 241
Db 139 PLTDPEVTNYSLRGC--EGKPLPKDLTFVADPKAGITIRNVKREYHRLCLHCSADQKGT 196

QY 242 -CTKLFTIDLNOAPOS-----TLPLQLFKVGEPLWIRCKAIHVNHGFLTWELDKALE 294
Db 197 VLSKKFTLKVRAIRAVPVVSVTSKLSLLEGEAFSVCFKOVSSSFVDSMW-----IK 250

QY 295 EGSYFEMSTYST-----NRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILE-- 347
Db 251 ENSQTNAQTOSNWHHGFDFERQEKLIISARVNDSGVFCMYANNFTFGSANVTITLEV 310

QY 348 -KGFINA-TSQOEYEDIDPEKFCFSVRFRAYPR-IRCTWIFSQASF--PCEQRGLEDEY 402
Db 311 DKGFIFNPFMMSTTIFNDGENVDLVIEYEAYPKPEHQQIYMMNRTETDKWEDYPKSDNE 370

QY 403 SISKEDCH-----KNKPGEYIFAENDDAQFTKMTLNLNRKPKQVLANASQA--SC 453
Db 371 SNRIYVELHLTRLKNGEGGYTFQVNSDVNSVVTENVVYNTKPEILLTHESLNGMLQC 430

QY 454 SSDGYPLPSTWTKKCDKSPNCTPEI-PEGVWNKANKRVFGQWVSSSTLNMSEAGKLL 512
Db 431 VVAGFPFPAVDWTFPCPAEQRCSPVIGPMDVQMGNSSLSFGKLVVQSSIDYSAFKHNGT 490

```

```

QY 513 VKCCAYNSMTSCETIFLNSPGPFPIQDN-----ISFYATIGLCLPFIIV 538
Db 491 VECRAYNWGRS--SAFNF-----FAFKGNSKEQIHPTTLFTPLLIGFVIAAGM---MCI 540

QY 559 LIVLCHIKYKKOFREYSQLOMI-QVTGPLDNEYFYDFRDYEDLKWEPFRENLEFGKVL 617
Db 541 IVMLITYKYLOKPMYEQWKVVEENG---NNVYIDPTQLPYDHWKPEPRNLSFGKTL 597

QY 618 GSGAFGRVWNTAYGISKTGVSQIQAIVAVKMLKEKADSCKEALMSELKMMTHLGHNDIVN 677
Db 598 GAGAFGKVVEATAYGLIKISDAAMTAVAKMLKPSAHLTEREALMSELKVLVSLGNHNMIVN 657

QY 678 LIGACTLSGPVYLIEFYCCYGDLLNLYRSKREKFHRTWTETFEKHNFPSSYPTQAHSNSS 737
Db 658 LIGACTVGPTTLVITEFYCCYGDLLNLYRKRKDSFCISKQEDHGEVALYKNLLSHKSSCS 717

QY 738 MPGRREVQLHPPLDQLSGFNGSIHSEDIENQKRLAEEDN-----VLTF 788
Db 718 DSTNEYMDMKP-----GDSYVVPYTKADKRRSRISYTERDVTYPAIMEDDELALDL 768

QY 789 EDLLCFAYQVAKMEFLEFKSCVHRDLAARNVLVTHGKVKICDFGLARDILSDSYVVR 848
Db 769 EDLLSFSYQVAKMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNVVK 828

QY 849 GNARLPVKWMAPESELEFGIYTIKSDVWSYGIILLWEIFSLGVNPPYGPVVDANFYKLIQSG 908
Db 829 GNARLPVKWMAPESEIFNCYITFESDVWSYGIIFLWELFSLGSSPYGPMVDSFYKMKIEG 888

QY 909 FKMEOPFATEGIYFVMOSWAFDSKRPSPFNLTSLFCQLAAEAEACIRTSIHLPKQA 968
Db 889 FRLMSPEHAPAEYDIKTCWDADPLKRPSTQIQLIEKQISDSTN-----HIYSNL 941

QY 969 APQQRGGLRAOSQPOQVKIHRER 991
Db 942 AN-----CSPNPERPVVDHVR 958

RESULT 13
Q98S02 ID Q98S02 PRELIMINARY; PRT; 977 AA.
AC Q98S02;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Receptor tyrosine kinase Fms.
OS Danio rerio (Kerr's danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=38750;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21359118; PubMed-11466528;
RA Parichy D.M., Johnson S.L.;
RT "Zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT diversification in Danio.";
RL Dev. Genes Evol. 211:319-328(2001).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF324480; AAK15302.1; -.
DR HSP; P11362; IFCK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00047; Ig; 4.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 4.

```


DR SMART; SM0040B; IGC2; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane.
FT VARIANT 31 P -> L.
FT VARIANT 148 L -> F.
FT VARIANT 218 R -> K.
FT VARIANT 263 I -> M.
FT VARIANT 490 R -> K.
FT VARIANT 594 E -> A.
FT VARIANT 614 G -> R.
FT VARIANT 926 N -> T.
FT VARIANT 934 A -> T.
FT VARIANT 938 A -> V.
SQ SEQUENCE 977 AA; 110294 MW; AAB2FF2A20D66090 CRC64;
Query Match 22.94; Score 1204.5; DB 13; Length 977;
Best Local Similarity 32.28; Pred. No. 5.9e-97;
Matches 334; Conservative 166; Mismatches 363; Indels 173; Gaps 36;
Qy 12 LLLVLSVILEFTVNODLPVKKVLISHENNGSSACKPSSRYMRVSGSPEDLOCTPRRQ 71
Db 5 LLLIGI---LFGQVGWSEPOIRL-----NSGAPAG--TDVILDGSGPQLVC----- 48
Qy 72 SEGTYEAATVEAEGSITITQVOLATPGDLSCLVFVHSHSLGCGPHFDLQNRGIVSMAI 131
Db 49 -EG-----DGPVTLPLRA-----KH-----KRYISKEVGKIRSRV 79
Qy 132 LNVETQAGEY-LHIIQSERANYVLTIVNVDTO-LVY---LRPYFRKMNQDALL-C 185
Db 80 EKATVDFGTGKVCYINENDSNLSNVHVFVRDRLVFLVPSITLRYVRK-EGEDLLPLC 138
Qy 186 ISEGVPEPTVENVLCSHRESCKEG-PANVRKEKVL----HELFGTDIRCCARNALGR 240
Db 139 L---LTDPDATDLMFRDMGSAAPYGMNATFDPKGLVLRNVHPGNADYICSAIRGAE 195
Qy 241 ECTKLFTDLNOAPOSTLPQLFLK-----VGEPLWIRCKAIHNVHGFGLTWELEDKAL 293
Db 196 KVKIFSINIOLRFP-PYVLRNEVYKLVGERLOISCTNNFNFFVYNTWTSHSRRL 254
Qy 294 EGSYFEMSTYSTNTMIRILLAFVSVGRNDTGYTCTSSSKHPQSALVT---ILEKGF 350
Db 255 PKAE--EKSTIEGRLAIESILT-IPSVQOSDNGNITCTGQNEAGANSSTTQLLVYDEPY 311
Qy 351 INAT---SSOEY-----EIDPYEKFCFSVRKAPRIRC---TWIFSQASFPCQR-- 396
Db 312 IRLSPKLSKLTIRGLSIEVSGDDVDGLVLEATPPPLTSHKWEPTTSINSLP-ENRY 370
Qy 397 GLEDGYISKFCDDHKN--KPGEYIFYAENDDAQFTKMTFLNIRKPKQVLAN-ASASQASC 453
Db 371 NHNDYEALLFLKRLNFEIEIGYTLNVNKMKSASITFDIKYTKPVARVKWENVTLSG 430
Qy 454 SSDGYPLPSWTWKKCDKSPNCTEIEI----PEGVNNKANRVKVGQWSSSTLNKSEAK 509
Db 431 RSYGYPAFSILWYQCTGRTTTPENTTDLPOTOTVFEQKESFGAVGESVLTVG-PNR 489
Qy 510 GLLVKCCAYNSMGTSCETIFLNSPGFFPIQDNISFYATIGLCLPFIV--VLIVLICHKY 567
Db 490 RMTVVYCAVFNLVGOGSDFTSMG-----VSDOIFTSAMCGSTVAMVGLGLLIFMIYK 542
Qy 568 KQRFYESOLOMIQVTPGLDNEYFYVDPRDYEVDLKWPEPPRENLEFGKVLGSGAFGRVNN 627
Db 543 KOKPRIERKWKIETATN--GNNTITDPTQLYNEKWEPPRKCLKGKTLGAGEFGKVVYE 600
Qy 628 ATAYGISKTGVSIOQVAVKMKADSCKEALMSSELKMMTHLGHHDNIYNLLGACATLSGP 687
Db 601 ATAYGLGKEDNITGVAVKMLKASAPDEREALMSSELKILSLHGQHKNIYNLLGACATHGCP 660

Qy 688 VYLIFEYCYGDLNLYLRKREKF-----HRTW-----TEI 718
Db 661 VLVITEYCHGDLNLYLRKREKF-----HRTW-----TEI 718
Qy 719 FKEHNFSSYPTTQAHSNNSMPGSRVOLHPPDLQSLGPNNGSIHSEDEIYEYENOKRLAE 778
Db 721 CSDHYLDMPVTSRPTNSTLDSSECO-----747
Qy 779 EEDLNVLTFFDLGCFAYOVAKGMEFEKSCVHRDLAARNVLTTHGKVVKICDFGLARD 838
Db 748 --EDSWPLDMDDLLRFSSQVAGLDFLAAKNCHIRDMAARNVLTNSRVAKICDFGLARD 805
Qy 839 ILSDSYVVRGNARLPVKWMAPESEFEGYITKSDVSYGILLWEIFSLGVNPPGIPVD 898
Db 806 INWDSYVVRGNARLPVKWMAPESEFEGYITKSDVSYGILLWEIFSLGVNPPGIPVD 865
Qy 899 ANFYKLIQSGFKMEOPFYATEGIYFVMQSCWAPDSKRSPNLTFLGCOLAE-AEEAC 957
Db 866 SKFYKMIKGYQMSRPDFAPEPMYTIKMCWNLDAAERTFSKISQIMQRLGETSEQD 925
Qy 958 IRTSIHLPKQAAPQOR 973
Db 926 NOEYKNIPAEAEQQ 941
RESULT 14
Q9TTD7
ID Q9TTD7 PRELIMINARY; PRT; 948 AA.
AC Q9TTD7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Stem cell factor receptor (Fragment).
GS C-KIT.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawrence S.B., Greenwood P., Tisdall D.J., McNatty K.P., Fidler A.E.;
RT "Partial cDNA sequence of the c-kit homologue of brushtail possum
RT (Trichosurus vulpecula).";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131209; RAF22141.1;
DR HSSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; ig; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00219; TyrcK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW Immunoglobulin domain; Receptor; Tyrosine-protein kinase.
FT NON_TER 948
SQ SEQUENCE 948 AA; 106323 MW; 708E0258FFC4D07D CRC64;
Query Match 22.88; Score 1201; DB 6; Length 948;
Best Local Similarity 33.08; Pred. No. 1.2e-96;
Matches 314; Conservative 148; Mismatches 356; Indels 134; Gaps 29;
Qy 77 YEATVEVAESGSIITQVOLATPGDLSC-----VWFKHSLLGCGPHFDLQNRGIVSMA 130
Db 33 YSSPSITPARS---QITVNFSGSEIKLSCSDLHFVNMVTFENAAAL-----HLE-SSRTTETWL 84
Qy 131 ILNVETQAGEYLLHIIQSERANYVLTIVNVDTO-LVY---LRPYFRKMNQDALL-C 185

Db 85 TNAKAODTGRYTC---TNKGLSSSIYFVKOPKILFLDLRLLYGNEGSDALVNCPTVD 141
QY 184 -----LCISEGVPEP-----TVEWLCSHRESCKEKGPAVVRKEEVLH 223
Db 142 PEVNTFLRCDGKPLPKDITLPIQOGKITKNK-RSHNKIC----- 184
QY 224 ELFGTDIRCCARNALGRECTKLEFTIDLNOAQPSTLPOL-----FLKVGPEPLWIRCKAI 276
Db 185 -----FQCSAYQDGLKLSDRMTLKVRAPOQS-VPEVSLLOTNVLREGFTQATCMIK 237
QY 277 HVNHGFLTWELEDKALEBESYEMSTYSTRNTRMIRILLAPVSVGRNDTGYTCSSSK- 335
Db 238 DVASSVVMWIKDNRRIT---STHTQSRHSGDYAYERQDILTISPVRVNDSGVFTCFANNT 294
QY 336 --HPSOSALVTILEKGFINA-TSQOEYEIDPYEKFCFSVRFKAYPR-IRCTWIFSQNSF 391
Db 295 FGANVTATLKVVEKGFINFPKMNNTIFINDGENIDLWVEYEAFFPEHQLQWYMGTV 354
QY 392 PCQORGLEDGYS-----ISKCDHNK-----PGEYIFVAENDDAQFTKMTLNI 436
Db 355 TDKW----DDYTRPGSESTIRYISEL--HLNRLKGTGGPYTFVSNSDVSTSVTFKVV 408
QY 437 RKQPVLANASAOA--SCSDGYPLSPWTKKCDKSPNC-TEEIPGVNKKANKRVF 493
Db 409 KTRPEILTSRLMNLQCVAAAGPEPTIDWYFCPGTEQRCSTSVLPMDVKNTSSMLPPF 468
QY 494 GQVSSSTLNKSEAGKGLLVKCCAYNSWGTSCETIFLNSPGPFFIQDNTISFYATIGLCL 553
Db 469 GKIVESTIDSSAFRYNGTVCKEASNDVGKS--SAFFN----FAIKQITSHLTFTPLLI 522
QY 554 PFIV-----VLIVLICHKKYKFRYEQSLQMI-QVTPGLDNEYFYDFRDEYDLKWEF 606
Db 523 GFVVAAGLMCVIILTYKYFQPMYEVQWVVEING---NNVYIDTQLPYDHKWEP 579
QY 607 PRENLEFGKVLGSGAFGRVNATAYGTSKTVSIOQAVKMLKEKADSCKEALMSLKMM 666
Db 580 PRNRLSFGKTLGAGAFGVVEATAYGLFKSDAAMTAVKMLKPSAHLTEREALMSLKVL 639
QY 667 THLGHDNIIVNLGACTLSGPVLIFFEYCCYDGLLNLVLRSKREKFRHTWTWIFKEH---- 722
Db 640 SYLGNHNIANLGACTTGGTIVTEYCCYDGLLNLFLRRKDSF-----ICSKHEDHA 693
QY 723 NFSYPTFQAHNSMPSGRE-VOLHPDLQSLGFGNSIHSDEIEVENOKRLAEDEEE 781
Db 694 EALYKNLQSKESSCDGANEDMDKPGVSVVPTKAEKRSARVGSYIERDVTATMED 753
QY 782 DLNLTEDLLCFAYQVAKMEFLEFKSVHRDLAARNVLVTHGKVKICDFGLARDILS 841
Db 754 DELALOTEDLLSFYSQVAKGMSFLASKNCIHRDLAARNILLTHGRITKICDFGLARHKN 813
QY 842 DSSVYVGNARLPVKWAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPGPVVDANF 901
Db 814 DSNVYVGNARLPVKWAPESIFNCVYTFSDVWSYGIWELFSLGSSPYPGMPVDSKF 873
QY 902 YKLIQSGFKMEQPFYATGEYFVVMQSCWAFDSRKRPSPNLTSLFGQLAE 953
Db 874 YKMIKGFRLSPPECAPPENMEINKSCWNEPDLQPTFKLIVLIEQQLLDS 925
RESULT 15
Q98SU4
ID Q98SU4 PRELIMINARY; PRT; 977 AA.
AC Q98SU4
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase Fms.
OS Danio albolineatus (pearl danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=27699;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
diversification in Danio";
RL Dev. Genes Evol. 211:319-328(2001).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
DR EMBL; AF324478; AAK15300.1; -.
DR HSSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTK_kinaseIII.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00047; ig; 4.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW phosphorylation; Receptor; Transmembrane.
FT VARIANT 57 57 L -> V.
FT VARIANT 164 164 A -> V.
FT VARIANT 325 325 R -> P.
FT VARIANT 326 326 G -> C.
FT VARIANT 327 327 L -> P.
FT VARIANT 328 328 S -> A.
FT VARIANT 328 328 S -> F.
FT VARIANT 366 366 E -> D.
FT VARIANT 398 398 Q -> K.
FT VARIANT 419 419 K -> R.
FT VARIANT 431 431 P -> R.
FT VARIANT 439 439 T -> S.
FT VARIANT 611 611 N -> I.
FT VARIANT 706 706 I -> T.
FT VARIANT 926 926 S -> T.
FT VARIANT 952 952 E -> D.
SQ SEQUENCE 977 AA; 109802 MW; FBF44AE73BDF463B CRC64;
Query Match 22.8%; Score 1198.5; DB 13; Length 977;
Best Local Similarity 33.3%; Pred. No. 2e-96;
Matches 325; Conservative 162; Mismatches 365; Indels 123; Gaps 33;
QY 84 VAESGSITLQVQLATPGDLSCLWFKHSLGCGPHFDLQNRGIV-----SMALNVTE 136
Db 37 IVDSGS---PLQLVCEGDAPVTFL-----PRLAKHKRYISKEVGKTRSFHVEKATV 84
QY 137 TQAGEY-LLHIQSERANYTVLFTVNVRTQ-LYV-----LRRPYERKMNQDALLCISEG 189
Db 85 DFTGTYKCVYINENDSNMSSSVHVFVRSRVLFVSPFSTSLR--YVRK-EGEDLLLPCLLT 141
QY 190 VPEPTVEWLCSHRESCKEKGPAVVRKEEVL-----HELFGTDIRCCARNALGRECTKL 245
Db 142 DPDAT-DFTFRMDNGSAAPYGMNATFDPKGVLRNHPGFNADYICSAIGCAEKVSKT 200
QY 246 FTIDLNOAQPSTLPOLFLK-----VGEPLWIRCKAIHVNHGFLTWELEDKALEEGSY 298
Db 201 FSNIIQRLRPP-PVYLKRNRYVVKLVGKELQISCTNNPNFSYNTVWTHSSRMLPKAE- 258
QY 299 FEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVT---ILEKGFNAT- 354

Db 259 -EKSTMEGDRLAIESILT-IPSVQSDTGNITCTGONEAGANSSTTQLLVVDEPYIRLSP 316
QY 355 --SSQEEY-----EIDPYKFCFSVRKAYPRIRC---TWIFSQASFPCEQR--GLEDG 401
Db 317 KLSSKLTHRLSTIESEGDDVDLGVLEIAYPLTSHKWTPTSHNASLP-ENRPNHNR 375
QY 402 YSISKFCDHKN--KPGEYIFYAENDDAQFTKMTFLAIRKKPQVLAN-ASASQASCSSDGY 458
Db 376 YEALLFLKRLNFEIGQYTLNVQNSMKNASITFDIKMYTKPAKVAKWENVYTLSCPSYGY 435
QY 459 PLPSWTWKCKSDKSPNCTEIEI----PEGVWNKANKRVFCQWVSSSTLNMSEAGKGLLVK 514
Db 436 PAPTILWYQCTGIRTCPCENTDLPQTQTVEFKESFGAVGVESVLTVG-PNRRMTVV 494
QY 515 CCAYNSMGTSCEIFLNSPFPFFIODNISFYATIGLCLPFIIV--VLIVLICHKKYKQER 572
Db 495 CVAFNLVGQSDTFSMD-----VSDQITSAMCGSTVAMVVLGGLLFIWIYKQKPR 547
QY 573 YESOLOMIOVTGPLDNEYFYVDFRDYEDYDLKWEPPRENLEFGKVLGSGAFGRVMNATAYG 632
Db 548 YEIRWKIIEATN--GNNYTFIDPTQLPYNERWEEPRDKLGLKTLGAGAFGKVVEATAYG 605
QY 633 ISKTGVSIOAVKMLKEKADSCKEALMSELKMMTHLGHHDNIIVNLLGACTLSGPVYLIF 692
Db 606 LGREDNITRVAVKMLKASAHDPDEREALMSELKILSHLQHKNIIVNLLGACTHGGPVLVIT 665
QY 693 EYCYGDLNLYLSKRKFKHR-----TWEIFKEHNFSSYPTFOAHSNNSMPGS----- 741
Db 666 EYCHGDLNLFRLSKAENFLNFVMTIPNFPMTDYKNVSIERMFVRSDSGISSTCSHY 725
QY 742 --REVOLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEEEEEDLNVLTFEDLLCFAYQV 798
Db 726 LDMRPVTSRTNSTLDS-----SSCEQEDSWPLDMDDDLRFSSQV 765
QY 799 AKGMFEFLKSCVHRDLAARNVLYTHGKVKIKCDFGLARDILSDSSYVYVGNARLPVKWM 858
Db 766 AQLDFLAARNCIHRDVAARNVLTNSRVAKICDFGLAPDIMNDSNYVYVGNARLPVKWM 825
QY 859 APESLPEGIYTIKSDVWSYGILLWEIFSLGYNYPGIPVDANFYKLIQSGFKMEQPPFYAT 918
Db 826 APESIFECVTVQSDVWSYGLWEIFSLGKSPYPNILLVDSKYFMKICGYOMSRPDPAP 885
QY 919 EGIYFVMSQWAFDSRKRPSPNLTSPFLGCOLAE-AEEACIRTSIHLPKQAAPQOQGLR 977
Db 886 PEMYIMKMCWNLDAERPTFSKISQMIQRLMGETSEQDSDQYKNIPIAEAGEQQ--LE 943
QY 978 AQSPQORQVKIHRERS 992
Db 944 SCDPAK-----HEEES 954

Search completed: May 27, 2003, 14:34:27
Job time : 83.4642 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 25, 2003, 02:59:08 ; Search time 8777.17 Seconds
(without alignments)
11608.404 Million cell updates/sec
Title: US-09-919-408-3
Perfect score: 3501
Sequence: 1 CGAGCGGCATCCGAGGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3501	100.0	3501	6	AR005212	AR005212 Sequence
2	3501	100.0	3501	6	AR071704	AR071704 Sequence
3	3501	100.0	3501	6	I25170	I25170 Sequence 3
4	3501	100.0	3501	6	I40601	I40601 Sequence 3
5	3475	99.3	3475	6	I44733	I44733 Sequence 3
6	3475	99.3	3475	6	HSU02687	U02687 Human growt
7	3474	99.2	3476	6	I40109	I40109 Sequence 1
8	3393.8	96.9	3422	9	HSFLT3RTPK	Z26652 H.sapiens F
9	3083.2	88.1	3120	6	I44518	I44518 Sequence 22
10	2274.6	65.0	3521	6	I44732	I44732 Sequence 1
11	2274.6	65.0	3521	10	MMFLT3	X59398 Mouse Flt3
12	2227.2	63.6	3453	6	AR005211	AR005211 Sequence
13	2227.2	63.6	3453	6	AR071703	AR071703 Sequence
14	2227.2	63.6	3453	6	I25169	I25169 Sequence 1
15	2227.2	63.6	3453	6	I40600	I40600 Sequence 1
16	2203	62.9	3426	10	MUSFLK2	M64689 Mouse flk-2
17	624.4	17.8	1894	6	AR149571	AR149571 Sequence
18	558.8	16.0	76721	9	AL591024	AL591024 Human DNA
19	476.8	13.6	480	9	HUM3RTPK	L36162 Homo sapien
20	400	11.4	400	11	G11126	G11126 human STS S
21	377.4	10.8	3545	5	AF153446	AF153446 Danio rer
22	336	9.6	3679	10	RRCSF1	X61479 Rat mRNA fo
23	330.4	9.4	2894	4	ECA224643	AJ224643 Equus cab
24	330.4	9.4	2973	4	ECA224642	AJ224642 Equus cab
25	330.2	9.4	4577	14	FCSSMONC	X01843 Feline sarc
26	329.8	9.4	3828	4	CATEFMC	J03149 Cat (P.dome
27	327.2	9.3	2894	4	ECA224644	AJ224644 Equus cab
28	326.6	9.3	429	10	MUS3RTPK	L36163 Mus musculu
29	325.6	9.3	2894	4	ECA224645	AJ224645 Equus cab
30	324.8	9.3	2924	4	AF131209	AF131209 Trichosur
31	324.6	9.3	2937	4	AF099030	AF099030 Canis fam
32	324.2	9.3	3992	6	AX394288	AX394288 Sequence
33	324.2	9.3	3992	9	HSCFMS	X03663 Human mRNA
34	323.2	9.2	2933	4	AF055037	AF055037 Equus cab
35	323.2	9.2	3575	10	MMCFCMS	X68932 M.musculus
36	323	9.2	2979	4	AF448148	AF448148 Canis fam.
37	322.6	9.2	2176	4	AF263826	AF263826 Bos tauru
38	322.6	9.2	2176	4	AF263827	AF263827 Bos tauru
39	322.6	9.2	3069	4	BOVCKR	DI6680 Bovine mRNA
40	322.2	9.2	3154	4	AF044249	AF044249 Canis fam
41	318.8	9.1	3826	5	XLXKRK1	Z48770 X.laavis C-
42	318.4	9.1	3655	10	MMEMSCR	X06368 Murine mRNA
43	317.2	9.1	3771	4	GOTCKIT	D45168 Goat mRNA f
44	317	9.1	5098	6	AX195906	AX195906 Sequence
45	317	9.1	5098	10	MMCKIT	Y00864 Mouse c-kit

ALIGNMENTS

RESULT 1
AR005212
LOCUS AR005212 3501 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5747651.
ACCESSION AR005212
VERSION AR005212.1 GI:3966091
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka,I.R.
TITLE Antibodies against tyrosine kinase receptor flk-1
JOURNAL Patent: US 5747651-A 3 05-MAY-1998;
FEATURES Location/Qualifiers


```
Db 2041 CTAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTGTAACCTGCTGGGGCGTGC 2100
QY 2101 ACATGTCAGGACCAATTTACTTTGATTTGAATACTGTGTGCTATGGTGATCTTCTCAAC 2160
Db 2101 ACATGTCAGGACCAATTTACTTTGATTTTGAATACTGTGTGCTATGGTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTACGTTTATACCCACATTTCAATCAGATCCAAATTCAGCATGCCGTGGTTCAAGA 2280
Db 2221 AATTTACGTTTATACCCACATTTCAATCAGATCCAAATTCAGCATGCCGTGGTTCAAGA 2280
QY 2281 GAAAGTTCAGATACACCGGACTCGCATCAAAATCTCAGGCTTCATGGGAATTCATTTCCAC 2340
Db 2281 GAAAGTTCAGATACACCGGACTCGCATCAAAATCTCAGGCTTCATGGGAATTCATTTCCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTGCTATCAAGTTGCCAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTGCTATCAAGTTGCCAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTACAGAGACCTGGCCGACGAAAGCTGCTTCCACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTGTTACAGAGACCTGGCCGACGAAAGCTGCTTCCACCCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATCCCAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATCCCAAC 2580
QY 2581 TATGTTGTCAGGGGCAATGCCCGTCTGCCGTGTAAATGGATGGCCCCCGAAGACCTGTTT 2640
Db 2581 TATGTTGTCAGGGGCAATGCCCGTCTGCCGTGTAAATGGATGGCCCCCGAAGACCTGTTT 2640
QY 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTATGGAATATGGAATATGGAATATGGAATAT 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTATGGAATATGGAATATGGAATATGGAATAT 2700
QY 2701 TTCTCAGTGTGTGAATCTTACCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAACTG 2760
Db 2701 TTCTCAGTGTGTGAATCTTACCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAACTG 2760
QY 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATA 2820
Db 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATA 2820
QY 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGGAACGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGGAACGCCATCTTCCCTAAATTTGACTTCG 2880
QY 2881 TTTTATGAGTATGAGTGGCAGATGAGAGAGGATGATATCAGAAATGATGGATGCGCT 2940
Db 2881 TTTTATGAGTATGAGTGGCAGATGAGAGAGGATGATATCAGAAATGATGGATGCGCT 2940
QY 2941 GTTTCGGAATGTCCTACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGGAATTTG 3000
Db 2941 GTTTCGGAATGTCCTACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGGAATTTG 3000
QY 3001 GGGTACTCTCTCCGAGGCTCAGTCTGGAAGATTCGTAGAGAACATTTAGTTTAAAG 3060
Db 3001 GGGTACTCTCTCCGAGGCTCAGTCTGGAAGATTCGTAGAGAACATTTAGTTTAAAG 3060
QY 3061 ACTTATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAGATTAATTTCAAT 3120
Db 3061 ACTTATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAGATTAATTTCAAT 3120
QY 3121 CACTAAAAGAAAATCTATTATCAACTGCTTCCACGACTTTTCTTAGAGGCGCTCT 3180
```

```
Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTTCCACGACTTTTCTTAGAGGCGCTCT 3180
QY 3181 GCGTTTACTCTTTGTTTTCAAAGGAGCTTTTGTAAAATCAAATCATCTGTCAAGAGCAG 3240
Db 3181 GCGTTTACTCTTTGTTTTCAAAGGAGCTTTTGTAAAATCAAATCATCTGTCAAGAGCAG 3240
QY 3241 GAGGAGCTGATATGAACCTTTATTTGAGGACTGATCTGATCCAAAGGCTTCTCAGGCGG 3300
Db 3241 GAGGAGCTGATATGAACCTTTATTTGAGGACTGATCTGATCCAAAGGCTTCTCAGGCGG 3300
QY 3301 GCTTGAGTCAATTTGCTACCTGAAGTACAGATATATTCTTTGTAATACATAAAAACAAAGC 3360
Db 3301 GCTTGAGTCAATTTGCTACCTGAAGTACAGATATATTCTTTGTAATACATAAAAACAAAGC 3360
QY 3361 ATTTTGGCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTA 3420
Db 3361 ATTTTGGCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTA 3420
QY 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTTCTACTACAGAAAA 3480
QY 3481 AAAAAAATAAAAAAATAAAAAA 3501
Db 3481 AAAAAAATAAAAAAATAAAAAA 3501

RESULT 2
LOCUS AR071704 3501 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5912133.
ACCESSION AR071704
VERSION AR071704.1 GI:7222592
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka, I. R.
TITLE Method for isolating stem cells expressing flk-1 receptors
JOURNAL Patent: US 5912133-A 3 15-JUN-1999;
FEATURES
    Location/Qualifiers
        source
            1..3501
            /organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGCGGATCCGAGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGCGGATCCGAGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGCGCTTGGCGCGGACCGCGGACCGCTGCCGCTGCTGTTTTTTCTGCAATGATA 120
Db 61 CCGCGCTTGGCGCGGACCGCGGACCGCTGCCGCTGCTGTTTTTTCTGCAATGATA 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCCGTGATCAAGTGTGTTTTTAATCATATAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCGTGATCAAGTGTGTTTTTAATCATATAAG 180
QY 181 AACAAATGATTCAGTGGGAAAGTATCATATATCCCATGTTATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTCAGTGGGAAAGTATCATATATCCCATGTTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGTGTGCGTTGAGACCCAGAGCTCAGGAGACGTGACGAAGCTGCCGCTGTG 300
Db 241 GACCTCGGTGTGCGTTGAGACCCAGAGCTCAGGAGACGTGACGAAGCTGCCGCTGTG 300
QY 301 GAAGTGGATGTATCTGCTTCATCACAAGTGTGCTGATGCCATGCCAGGAAACATT 360
Db 301 GAAGTGGATGTATCTGCTTCATCACAAGTGTGCTGATGCCATGCCAGGAAACATT 360
```

Db 301 GAAAGTGATGTATCTGCTCCATCACACTGCAAGTCTGGTCGATGCCCCAGGAACATT 360
Qy 361 TCCTGCTCTGGGTCTTTAAGCACAGCTCCCTGAATGGCCAGCCACATTTTCATTACAA 420
Db 361 TCCTGCTCTGGGTCTTTAAGCACAGCTCCCTGAATGGCCAGCCACATTTTCATTACAA 420
Qy 421 AACAGAGAGTGTTCATGCTCATTTTGAATAAGACAGAAACCCAGCTGGAGAATAC 480
Db 421 AACAGAGAGTGTTCATGCTCATTTTGAATAAGACAGAAACCCAGCTGGAGAATAC 480
Qy 481 CTACCTTTTATTCAGAGTGAAGCTACCAATACAAATATGTTTACAGTGAATTAAGA 540
Db 481 CTACCTTTTATTCAGAGTGAAGCTACCAATACAAATATGTTTACAGTGAATTAAGA 540
Qy 541 AATACCCCTGCTTTACACATTAGAAGACCTTACTTTAGAAAATGGAACACAGCAGCC 600
Db 541 AATACCCCTGCTTTACACATTAGAAGACCTTACTTTAGAAAATGGAACACAGCAGCC 600
Qy 601 CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGCTTTGCCGATTCA 660
Db 601 CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGCTTTGCCGATTCA 660
Qy 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAGTGCTT 720
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAGTGCTT 720
Qy 721 CATGAATTTATTTGGACGGACATAGGTGCTGTGCCAGAAATCAACTGGCGAGGAATGC 780
Db 721 CATGAATTTATTTGGACGGACATAGGTGCTGTGCCAGAAATCAACTGGCGAGGAATGC 780
Qy 781 ACCAGGCTGTTCCACAAATAGATCTAAATCAAACTCTCGACACACATTTGCCACAATATT 840
Db 781 ACCAGGCTGTTCCACAAATAGATCTAAATCAAACTCTCGACACACATTTGCCACAATATT 840
Qy 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCTGTTGCTTTGCTATCATCAGTGGA 900
Db 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCTGTTGCTTTGCTATCATCAGTGGA 900
Qy 901 GGGCTCACTGGGAATTAGAAACAAAGCCTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Db 901 GGGCTCACTGGGAATTAGAAACAAAGCCTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGCTATCATCAGTGGA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGCTATCATCAGTGGA 1020
Qy 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCGTAGGAAGGGATTATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140
Db 1081 GTTACCATCGTAGGAAGGGATTATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140
Qy 1141 GACCAATATGAAGTGTGTTTCTGTCAGTTTAAAGCCTTACCACCAAAATCAGATGT 1200
Db 1141 GACCAATATGAAGTGTGTTTCTGTCAGTTTAAAGCCTTACCACCAAAATCAGATGT 1200
Qy 1201 ACGTGGACCTCTCTCGAAAATCATTTCCCTGTGTGAGCAAAAGGCTTGTATACGGATAC 1260
Db 1201 ACGTGGACCTCTCTCGAAAATCATTTCCCTGTGTGAGCAAAAGGCTTGTATACGGATAC 1260
Qy 1261 AGCATATCAAGTTTTCGAATCATAGCACAGCCAGGAGAAATATATTTCCATGCGAGAA 1320
Db 1261 AGCATATCAAGTTTTCGAATCATAGCACAGCCAGGAGAAATATATTTCCATGCGAGAA 1320
Qy 1321 AATGATGATGCCAAATTTACCRAAAATGTCACGCTGAATATAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAAATTTACCRAAAATGTCACGCTGAATATAGAGGAAACCTCAAGTG 1380
Qy 1381 CTCGAGAGAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTACCATCT 1440
Db 1381 CTCGAGAGAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTACCATCT 1440

Qy 1441 TGGACTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGAGATCACAGAAGGA 1500
Db 1441 TGGACTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGAGATCACAGAAGGA 1500
Qy 1501 GTCTGGAATPAGAAGGCTAACAGAAAAGTGTGAGAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATPAGAAGGCTAACAGAAAAGTGTGAGAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 ACATGAGTGAAGCCATAAAAGGTTCCCTGAGTCAAGTGTGTCATACAAATCCCTTGGC 1620
Db 1561 ACATGAGTGAAGCCATAAAAGGTTCCCTGAGTCAAGTGTGTCATACAAATCCCTTGGC 1620
Qy 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680
Db 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680
Qy 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTTCAATGTCGTTTTTAAACCTTGCTA 1740
Db 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTTCAATGTCGTTTTTAAACCTTGCTA 1740
Qy 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCTCAGATAAATGAGTACTTCTACGTTGATTTTTCAGAGAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAAATGAGTACTTCTACGTTGATTTTTCAGAGAATATGAATATGATCTC 1860
Qy 1861 AAATGGAGTGTTCAGAGAAAATTTAGAGTTTGGGAAGGCTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGAGTGTTCAGAGAAAATTTAGAGTTTGGGAAGGCTACTAGGATCAGGTGCTTTT 1920
Qy 1921 GGAAGAGTATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGTT 1980
Db 1921 GGAAGAGTATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGTT 1980
Qy 1981 GCGCTCAAAATGCTGAAAAGAAAGCAGACAGCTCTGAAAAGAGGACACTCATGTTCAGAA 2040
Db 1981 GCGCTCAAAATGCTGAAAAGAAAGCAGACAGCTCTGAAAAGAGGACACTCATGTTCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCTGC 2100
Qy 2101 ACATGTCAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTCTCAAC 2160
Db 2101 ACATGTCAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAAGAGAAAATTTTACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAAGAGAAAATTTTACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTTCAGTTTTTACCCCACTTTTCCAAATCATCTCAAAATTTCCAGCATGCTGGTTCAAGA 2280
Db 2221 AATTTTCAGTTTTTACCCCACTTTTCCAAATCATCTCAAAATTTCCAGCATGCTGGTTCAAGA 2280
Qy 2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTTCAAC 2340
Db 2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTTCAAC 2340
Qy 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTTGAAGATCTTCTTTGCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTTGAAGATCTTCTTTGCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTTGCCCGCCAGGACGCTGCTTGTACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTTGCCCGCCAGGACGCTGCTTGTACCCAC 2520


```
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db 2581 TATGTTGTGAGGGCAATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy 2641 GAAGGCTCTACACCATTAAGAGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 GAAGGCTCTACACCATTAAGAGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy 2701 TTCTCACTTGGTGTGAATCCTTACCTGCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Db 2701 TTCTCACTTGGTGTGAATCCTTACCTGCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Qy 2761 ATTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 ATTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Qy 2821 ATGCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Db 2821 ATGCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Qy 2881 TTTTGTAGGATGTCAGCTGGCAGATGCGAAGAGCGATGATCAGAATGTGGATGGCGCT 2940
Db 2881 TTTTGTAGGATGTCAGCTGGCAGATGCGAAGAGCGATGATCAGAATGTGGATGGCGCT 2940
Qy 2941 GTTTCGGAATGCTCCTACACCTACCAAACAGGCGACCTTTTCAGCAGAGAGATGGATTTG 3000
Db 2941 GTTTCGGAATGCTCCTACACCTACCAAACAGGCGACCTTTTCAGCAGAGAGATGGATTTG 3000
Qy 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATGCTAGAGAGCAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATGCTAGAGAGCAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCTAACAGGCTGTAGATTACCAACAAAGATTAATTTTCAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCTAACAGGCTGTAGATTACCAACAAAGATTAATTTTCAT 3120
Qy 3121 CACTTAAAGAAATCTATATCAACTGCTGCTTCAACAGACTTTTCTCTAGAACCGCTCT 3180
Db 3121 CACTTAAAGAAATCTATATCAACTGCTGCTTCAACAGACTTTTCTCTAGAACCGCTCT 3180
Qy 3181 GCGTTTACTCTGTTTCAAGGACCTTTTGTAAATCAATATCATCTGTCACAGGCGAG 3240
Db 3181 GCGTTTACTCTGTTTCAAGGACCTTTTGTAAATCAATATCATCTGTCACAGGCGAG 3240
Qy 3241 GAGGAGCTGATAATGAAGACTTTTATGGAGCATGATCTGCATCCAAGGCTTCTCAGGCGC 3300
Db 3241 GAGGAGCTGATAATGAAGACTTTTATGGAGCATGATCTGCATCCAAGGCTTCTCAGGCGC 3300
Qy 3301 GCTTGTAGTGAATGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
Db 3301 GCTTGTAGTGAATGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
Qy 3361 ATTTTGTAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTHAA 3420
Db 3361 ATTTTGTAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTHAA 3420
Qy 3421 TTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
```

RESULT 3
125170
LOCUS
DEFINITION Sequence 3 from patent US 5548065.
125170 3501 bp DNA linear PAT 07-OCT-1996

```
ACCESSION I25170
VERSION I25170.1 GI:1605040
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3501)
TITLE Lemischka, I.R.
JOURNAL Tyrosine kinase receptor human flk-2-specific antibodies
FEATURES Patent: US 5548065-A 3 20-AUG-1996;
Location/Qualifiers
1..3501
source /organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCCCTGGGGGACCCGGGCTCCGGAGGCGCATG 60
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCCCTGGGGGACCCGGGCTCCGGAGGCGCATG 60
Qy 61 CGCGGTTGGGCGCGACGCGGCGCACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 CGCGGTTGGGCGCGCGACGCGGCGCACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCCGTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCGTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180
Qy 181 AACATGATTCATAGTGGGAGTCAATCATATATCCATGGTATCAGAAATCCCGGAA 240
Db 181 AACATGATTCATAGTGGGAGTCAATCATATATCCATGGTATCAGAAATCCCGGAA 240
Qy 241 GACCTCGGCTGCTGAGAGCCGAGAGCTCAGGACAGTGTACGAACTGCGGCTGTG 300
Db 241 GACCTCGGCTGCTGAGAGCCGAGAGCTCAGGACAGTGTACGAACTGCGGCTGTG 300
Qy 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 361 TCCTGCTCTGCTGCTTTAAGACAGCTCCCTGAATGCGCAGCAGCATTTGATTACAA 420
Db 361 TCCTGCTCTGCTGCTTTAAGACAGCTCCCTGAATGCGCAGCAGCATTTGATTACAA 420
Qy 421 AACAGAGAGTGTGTTTCCATGCTATTTGAAATGACAGAAACCCCAAGCTGGAGAAATAC 480
Db 421 AACAGAGAGTGTGTTTCCATGCTATTTGAAATGACAGAAACCCCAAGCTGGAGAAATAC 480
Qy 481 CTACTTTTATTACAGAGTGAAGTACCATTACAAATATTGTTTACAGTGAATATAAGA 540
Db 481 CTACTTTTATTACAGAGTGAAGTACCATTACAAATATTGTTTACAGTGAATATAAGA 540
Qy 541 AATACCCCTGCTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGCGCC 600
Db 541 AATACCCCTGCTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGCGCC 600
Qy 601 CTGGTCTGCATATCTGAGAGCGCTTCCAGAGCGGATCGTGGAAATGCGTGTGCTGCGATTCA 660
Db 601 CTGGTCTGCATATCTGAGAGCGCTTCCAGAGCGGATCGTGGAAATGCGTGTGCTGCGATTCA 660
Qy 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTGCCT 720
Db 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTGCCT 720
Qy 721 CATGAATTTATTTGGGACGACATAAGTGTGCTGCGAGAAATGAATCGGCGAGGGAATGC 780
Db 721 CATGAATTTATTTGGGACGACATAAGTGTGCTGCGAGAAATGAATCGGCGAGGGAATGC 780
Qy 781 ACCAGGCTGTTCAACAATAGATCTAAATCAACTCTCAGACCACATTTGCCACAATATTTT 840
```

Db 781 ACCAGGCTGTTCCACAAATAGATCTAAATCAAACTCCTCAGACACAAATGTCACAAATATTT 840
Qy
Db 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCTATGTTGAACCATGGATTC 900
Db 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCTATGTTGAACCATGGATTC 900
Qy 901 GGGCTCAGCTGGGAATTAAGAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Db 901 GGGCTCAGCTGGGAATTAAGAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTGTTATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTGTTATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGACATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGACATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTATGAAAT 1140
Db 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTATGAAAT 1140
Qy 1141 GACCAATATGAAAGATTTGTTTCTGTACAGTTTAAAGCTACCCACAAATCAGATGT 1200
Db 1141 GACCAATATGAAAGATTTGTTTCTGTACAGTTTAAAGCTACCCACAAATCAGATGT 1200
Qy 1201 AGCTGGACCTCTCTCGAAAATCATTTCCCTGTGAGCAAAAGGGTCTTGTATACGGATAC 1260
Db 1201 AGCTGGACCTCTCTCGAAAATCATTTCCCTGTGAGCAAAAGGGTCTTGTATACGGATAC 1260
Qy 1261 AGCATATCCAAGTTTGGCAATCATAAAGCACCGCAGGAGAAATATATATTCATCGAGAA 1320
Db 1261 AGCATATCCAAGTTTGGCAATCATAAAGCACCGCAGGAGAAATATATATTCATCGAGAA 1320
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGTAATATAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGTAATATAGAGGAAACCTCAAGTG 1380
Qy 1381 CTCGAGAAAGCATCGGCAAGTCAGCGCTCTGTTCTCGATGGATACCCATTACCATCT 1440
Db 1381 CTCGAGAAAGCATCGGCAAGTCAGCGCTCTGTTCTCGATGGATACCCATTACCATCT 1440
Qy 1441 TGGACCTGGAAGAAGTGTTCAGCAAGTCTCCAACTGCACAGAGAGATCAGAGAAG 1500
Db 1441 TGGACCTGGAAGAAGTGTTCAGCAAGTCTCCAACTGCACAGAGAGATCAGAGAAG 1500
Qy 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGTTGACAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGTTGACAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATAAAGGGTTCGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Qy 1621 ACATCTTGTGAGACGATCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680
Db 1621 ACATCTTGTGAGACGATCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680
Qy 1681 ATCTCATTCTATGCAACAATTTGGTGTGTTCTCTCTCTTCTCATTGTCGTTTAAACCTGCTA 1740
Db 1681 ATCTCATTCTATGCAACAATTTGGTGTGTTCTCTCTCTTCTCATTGTCGTTTAAACCTGCTA 1740
Qy 1741 ATTTGTACAAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTACAAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGCTCCTCAGATAATCAGTACTTCTACGTTGATTTTCAGAGAATATCAATATGATCTC 1860
Db 1801 ACCGCTCCTCAGATAATCAGTACTTCTACGTTGATTTTCAGAGAATATCAATATGATCTC 1860
Qy 1861 AAATGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Db

Db 1861 AAATGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAAAGGTACTAGATCAGGTGCTTTT 1920
Qy 1921 GGAAAGTGTATGAACGAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGTTT 1980
Db 1921 GGAAAGTGTATGAACGAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGTTT 1980
Qy 1981 GCCTGCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040
Db 1981 GCCTGCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGGAAATTTGCTGAACCTGCTGGGGGCTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGGAAATTTGCTGAACCTGCTGGGGGCTGC 2100
Qy 2101 ACACGTCTCAGGACCAATTTACTTGTGTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160
Db 2101 ACACGTCTCAGGACCAATTTACTTGTGTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAAAGTAAAGAGAAAATTTTCAGAGACTTGGACAGAGATTTTCAGAGGAACAC 2220
Db 2161 TATCTAAGAAAGTAAAGAGAAAATTTTCAGAGACTTGGACAGAGATTTTCAGAGGAACAC 2220
Qy 2221 AATTTTCAGTTTACCCCACTTCCCAATCAGATCCTCAAAATTTCCAGCATGCTGTTCAAGA 2280
Db 2221 AATTTTCAGTTTACCCCACTTCCCAATCAGATCCTCAAAATTTCCAGCATGCTGTTCAAGA 2280
Qy 2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340
Db 2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATGATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATGATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAAGATCTTCTTCTGCTTGTGATATCAAGTTCGCAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAAGATCTTCTTCTGCTTGTGATATCAAGTTCGCAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCGTGTTCACAGAGACCTCGCCGCCAGGAAACGTCGTTCACCCAC 2520
Db 2461 CTGGAATTTAAGTCGTGTTCACAGAGACCTCGCCGCCAGGAAACGTCGTTCACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACAC 2580
Qy 2581 TATGTTGTCAGGGGCAATGCCGCTGTCCTGTAATGGAATGGATGGCCCCGAAAGCCTGTTT 2640
Db 2581 TATGTTGTCAGGGGCAATGCCGCTGTCCTGTAATGGAATGGATGGCCCCGAAAGCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700
Db 2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700
Qy 2701 TTTCTACTTGGTGAATCCTTACCCTGGCATTTCGGTTGATGCTAACTTCTTCAAACTG 2760
Db 2701 TTTCTACTTGGTGAATCCTTACCCTGGCATTTCGGTTGATGCTAACTTCTTCAAACTG 2760
Qy 2761 ATTCAAAATGGATTTAAAATGGATGATGCCCATTTTATGCTTACAGAGAAATATACATTATA 2820
Db 2761 ATTCAAAATGGATTTAAAATGGATGATGCCCATTTTATGCTTACAGAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAAGCCATCTTCCCTAATTTGACCTTCG 2880
Db 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAAGCCATCTTCCCTAATTTGACCTTCG 2880
Qy 2881 TTTTATAGGATGTGAGTGGCAGATGCAAGAAAGCAGTATGATCAGAAATGAGTGGATGCGCT 2940
Db 2881 TTTTATAGGATGTGAGTGGCAGATGCAAGAAAGCAGTATGATCAGAAATGAGTGGATGCGCT 2940
Qy 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTG 3000
Db 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTG 3000

```
QY 3001 GGGCTACTCTCTCCGAGGCTCAGTCTGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
|||||
Db 3001 GGGCTACTCTCTCCGAGGCTCAGTCTGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT 3120
|||||
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT 3120
QY 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAACCGGCT 3180
|||||
Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAACCGGCT 3180
QY 3181 CGGTTTACTCTGTTTCCAAAGGACTTTTGTAAATCAAAATCATCTGTCACAGGGCAG 3240
|||||
Db 3181 CGGTTTACTCTGTTTCCAAAGGACTTTTGTAAATCAAAATCATCTGTCACAGGGCAG 3240
QY 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCAATGATCTGCATCCAAAGGCTTCTCAGGCGG 3300
|||||
Db 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCAATGATCTGCATCCAAAGGCTTCTCAGGCGG 3300
QY 3301 GCTTGAGTGAATTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3360
|||||
Db 3301 GCTTGAGTGAATTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3360
QY 3361 ATTTTGTAAAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATGTAAA 3420
|||||
Db 3361 ATTTTGTAAAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATGTAAA 3420
QY 3421 TTTTTCAGTATTAGTATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
|||||
Db 3421 TTTTTCAGTATTAGTATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 4
140601
LOCUS 3501 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 3 from patent US 5621090.
ACCESSION I40601
VERSION I40601.1 GI:2082893
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka,I.R.
TITLE Nucleic acids encoding soluble human FLK-2 extracellular domain
JOURNAL Patent: US 5621090-A 3 15-APR-1997;
FEATURES
Location/Qualifiers
1..3501
/organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGCGGCATCCGAGGCTGGCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
|||||
Db 1 CGAGGCGGCATCCGAGGCTGGCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGGCTTGGCGGCGGAGCGGGACCGCTGCCGCTCGCTGCTGTTTCTGCAATGATA 120
|||||
Db 61 CCGGCTTGGCGGCGGAGCGGGACCGCTGCCGCTCGCTGCTGTTTCTGCAATGATA 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
|||||
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
```

```
QY 181 AACAATGATTTCATCAGTGGGAAGTCATCATATATCCCATGTATCAGAAATCCCGGAA 240
|||||
Db 181 AACAATGATTTCATCAGTGGGAAGTCATCATATATCCCATGTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGGTGTGCGGTTGAGAGCCCAAGAGCTCAGGACAGCTGTACGAAGCTGCGGCTGTG 300
|||||
Db 241 GACCTCGGGTGTGCGGTTGAGAGCCCAAGAGCTCAGGACAGCTGTACGAAGCTGCGGCTGTG 300
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTCTGCTGCTGATGCCAGGAACATTT 360
|||||
Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTCTGCTGCTGATGCCAGGAACATTT 360
QY 361 TCCTGCTCTCTGCTGCTTTAAGCAGACGCTCCCTCAATTTGCCAGCCACATTTTGTATTTACAA 420
|||||
Db 361 TCCTGCTCTCTGCTGCTTTAAGCAGACGCTCCCTCAATTTGCCAGCCACATTTTGTATTTACAA 420
QY 421 AACAGAGGAGTGTGTTTCCATGTCTATTGAAATGACAGAAACCCCAAGCTGGAGATAC 480
|||||
Db 421 AACAGAGGAGTGTGTTTCCATGTCTATTGAAATGACAGAAACCCCAAGCTGGAGATAC 480
QY 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAATATAAGA 540
|||||
Db 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAATATAAGA 540
QY 541 AATACCTGCTTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACAGGAGGCC 600
|||||
Db 541 AATACCTGCTTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACAGGAGGCC 600
QY 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGAATGGTGTCTTCGCGATTCA 660
|||||
Db 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGAATGGTGTCTTCGCGATTCA 660
QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
|||||
Db 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
QY 721 CATGAATATTGTTGGACGACATAGGTGCTGTGCCAGAAATGAATGCGCAGGGAATGC 780
|||||
Db 721 CATGAATATTGTTGGACGACATAGGTGCTGTGCCAGAAATGAATGCGCAGGGAATGC 780
QY 781 ACCAGGCTCTTCACAAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAAATATTT 840
|||||
Db 781 ACCAGGCTCTTCACAAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAAATATTT 840
QY 841 CTTAAAGTAGGGGAACCCCTTATGATAGTGCAGAGCTGTTTCATGTGAACCATGGATTC 900
|||||
Db 841 CTTAAAGTAGGGGAACCCCTTATGATAGTGCAGAGCTGTTTCATGTGAACCATGGATTC 900
QY 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGAGGGCAACTACTTTGAGATGAGT 960
|||||
Db 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGAGGGCAACTACTTTGAGATGAGT 960
QY 961 ACCTATTCAACAAACAGAACTATGATACGATTCTCTTTGCTTTTATCATCAGTGGCA 1020
|||||
Db 961 ACCTATTCAACAAACAGAACTATGATACGATTCTCTTTGCTTTTATCATCAGTGGCA 1020
QY 1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
|||||
Db 1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCGTAGGAAAGGATTTATAATGTCTACAAATCAAGTGAAGATTTGAAATTT 1140
|||||
Db 1081 GTTACCATCGTAGGAAAGGATTTATAATGTCTACAAATCAAGTGAAGATTTGAAATTT 1140
QY 1141 GACCAATATGAGAGTGTGTTTCTGTGAGTGTGTTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 1200
|||||
Db 1141 GACCAATATGAGAGTGTGTTTCTGTGAGTGTGTTTAAAGCTTAAAGCTTAAAGCTT 1200
QY 1201 AGCTGGAGCTTCTCTCGAAAAATCATTTTCTGTGAGCAAAAGGCTCTTGATAACGATAC 1260
|||||
Db 1201 AGCTGGAGCTTCTCTCGAAAAATCATTTTCTGTGAGCAAAAGGCTCTTGATAACGATAC 1260
```

QY 1261 AGCATATCCAAAGTTTTCATATCAATCAAGACACAGCCAGGAGAATATATATTCATGCGAGAA 1320
DB 1261 AGCATATCCAAAGTTTTCATATCAATCAAGACACAGCCAGGAGAATATATATTCATGCGAGAA 1320
QY 1321 AATGATGATGCCAAATTTACCAAAATGTTTACGCTGTAATATAAGAAAGAAACCTCAAGTG 1380
DB 1321 AATGATGATGCCAAATTTACCAAAATGTTTACGCTGTAATATAAGAAAGAAACCTCAAGTG 1380
QY 1381 CTCCGAGAAGCATCGGCAAGTCAGCGCTCCCTGTTTCTCGGATGGATACCCATTACCATCT 1440
DB 1381 CTCCGAGAAGCATCGGCAAGTCAGCGCTCCCTGTTTCTCGGATGGATACCCATTACCATCT 1440
QY 1441 TGGACCTTGGAAAGAGTGTTCAGACAAGTCTCCCAACTGTCAGACAAGAGATCACAGAAGGA 1500
DB 1441 TGGACCTTGGAAAGAGTGTTCAGACAAGTCTCCCAACTGTCAGACAAGAGATCACAGAAGGA 1500
QY 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
DB 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCGTCAGTGTCTGTCATACAAATTCCTTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCGTCAGTGTCTGTCATACAAATTCCTTTGGC 1620
QY 1621 ACATCTTGTGAGACGATCCCTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
DB 1621 ACATCTTGTGAGACGATCCCTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTTCTCCCTTCATGTCGTTTAACTCTGCTA 1740
DB 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTTCTCCCTTCATGTCGTTTAACTCTGCTA 1740
QY 1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
QY 1801 ACCGCTCCTCAGATAATGAGTACTTCTAGCTTGTGTTTGGGAGGTTAGGATCAGGTGCTTTT 1920
DB 1801 ACCGCTCCTCAGATAATGAGTACTTCTAGCTTGTGTTTGGGAGGTTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGTGTATGAAAGCAACCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
DB 1921 GGAAGTGTATGAAAGCAACCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
QY 1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
DB 1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
QY 2041 CTCAGATGATGACCCAGCTGGGAGCCACAGAGATATTTGTAACCTGCTGGGGGCTGC 2100
DB 2041 CTCAGATGATGACCCAGCTGGGAGCCACAGAGATATTTGTAACCTGCTGGGGGCTGC 2100
QY 2101 ACACCTCAGGACCAATTTACTTTGATTTTGTGATCTGTTGATCTGTTGATCTGTTCAAC 2160
DB 2101 ACACCTCAGGACCAATTTACTTTGATTTTGTGATCTGTTGATCTGTTGATCTGTTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGCTTTGGACAGAGATATTTTCAGGAAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGCTTTGGACAGAGATATTTTCAGGAAACAC 2220
QY 2221 AATTTCAAGTTTACCCCACTTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
DB 2221 AATTTCAAGTTTACCCCACTTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
QY 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGATTCATTTCCAC 2340
DB 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGATTCATTTCCAC 2340
QY 2341 TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400

DB 2341 TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTTTGTCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTTTGTCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCACAGACACCTGGCCGCGCAGGAACGCTGCTGTCAACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCACAGACACCTGGCCGCGCAGGAACGCTGCTGTCAACCCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATATCATGAGTGAATTCAC 2580
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATATCATGAGTGAATTCAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGGATGGCCCCGGAAGACCTGTTT 2640
DB 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGGATGGCCCCGGAAGACCTGTTT 2640
QY 2641 GAAGGCATCTACACCATTAAGAGTGTCTGCTGATATGGAATATTTACTGTGGGAAATC 2700
DB 2641 GAAGGCATCTACACCATTAAGAGTGTCTGCTGATATGGAATATTTACTGTGGGAAATC 2700
QY 2701 TTCTCAGTTGTTGTAATCTTACCTGGCATTCGGGTTGATGCTAACTTCTACAAACTG 2760
DB 2701 TTCTCAGTTGTTGTAATCTTACCTGGCATTCGGGTTGATGCTAACTTCTACAAACTG 2760
QY 2761 ATTTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTTACAGAGAATAATACATATA 2820
DB 2761 ATTTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTTACAGAGAATAATACATATA 2820
QY 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGAAACGGCCATCTTCCCTTAATTTGACTTGC 2880
DB 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGAAACGGCCATCTTCCCTTAATTTGACTTGC 2880
QY 2881 TTTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2940
DB 2881 TTTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2940
QY 2941 GTTTCGGAAATGCTTACACCTTACCAAAACAGGCGACCTTTTACAGAGAGATGGAATTTG 3000
DB 2941 GTTTCGGAAATGCTTACACCTTACCAAAACAGGCGACCTTTTACAGAGAGATGGAATTTG 3000
QY 3001 GGGCTACTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
DB 3001 GGGCTACTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCTCCCTATCCCTTAACAGGCTGTAGATTACCAAAACAGGAATTAATTTTCAAT 3120
DB 3061 ACTTCATCTCCCTATCCCTTAACAGGCTGTAGATTACCAAAACAGGAATTAATTTTCAAT 3120
QY 3121 CACTAAAGAAAATCTATTATCAACTGCTTTCACAGACTTTTCTTCTAGAGCCGCTCT 3180
DB 3121 CACTAAAGAAAATCTATTATCAACTGCTTTCACAGACTTTTCTTCTAGAGCCGCTCT 3180
QY 3181 GGGTTTACTCTTGTTCGAGGCTTTTGTGTAATCAATCATCTGTCACAGGAG 3240
DB 3181 GGGTTTACTCTTGTTCGAGGCTTTTGTGTAATCAATCATCTGTCACAGGAG 3240
QY 3241 GAGGAGCTGATATGAACTTTATTTGGAGCATTCATCTGCAAGGCTTCTCAGGCG 3300
DB 3241 GAGGAGCTGATATGAACTTTATTTGGAGCATTCATCTGCAAGGCTTCTCAGGCG 3300
QY 3301 GCTTGAAGTAAATGTGTACCTGAAGTACAGTATTTCTTGAATACATAAAACAAAGC 3360
DB 3301 GCTTGAAGTAAATGTGTACCTGAAGTACAGTATTTCTTGAATACATAAAACAAAGC 3360
QY 3361 ATTTTGTAAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
DB 3361 ATTTTGTAAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
QY 3421 TTTTTCAGCTATTTAGTGTATATTTTATGGTGGGGAATTAATTTCTACTACAGAAA 3480

Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480

Qy 3481 AAAAAAAAAAAAAAAAAAAAA 3501

Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

RESULT 5

LOCUS I44733 3475 bp DNA linear PAT 07-OCT-1997

DEFINITION Sequence 3 from patent US 5635388.

ACCESSION I44733

VERSION I44733.1 GI:2469446

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3475)

AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.

TITLE Agonist antibodies against the flk2/flt3 receptor and uses thereof

JOURNAL Patent: US 5635388-A 3 03-JUN-1997;

FEATURES

Location/Qualifiers

1..3475

/organism="unknown"

BASE COUNT 1042 a 709 c 784 g 940 t

ORIGIN

Query Match 99.3%; Score 3475; DB 6; Length 3475;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCAGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

Db 1 CGAGCGGCATCCAGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

Qy 61 CCGCGCTTGGCGCGCAGCGCGGCACCGTGGCTGCTGTTTCTTTCGCAATGATA 120

Db 61 CCGCGCTTGGCGCGCAGCGCGGCACCGTGGCTGCTGTTTCTTTCGCAATGATA 120

Qy 121 TTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Db 121 TTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Qy 181 AACATGATTCATCAGTGGGAAGTCATCATATCCATGCTGATCAGATCCCGGAA 240

Db 181 AACATGATTCATCAGTGGGAAGTCATCATATCCATGCTGATCAGATCCCGGAA 240

Qy 241 GACCTCGGGTGTGGTGTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300

Db 241 GACCTCGGGTGTGGTGTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300

Qy 301 GAAGTGGATGTATCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Db 301 GAAGTGGATGTATCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 361 TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAAATGCCAGCCACATTTGATTTACAA 420

Db 361 TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAAATGCCAGCCACATTTGATTTACAA 420

Qy 421 AACAGAGAGTGTGTTTCCATGCTTGTGAAATGACAGAAACCCAGCTGGAGAATAC 480

Db 421 AACAGAGAGTGTGTTTCCATGCTTGTGAAATGACAGAAACCCAGCTGGAGAATAC 480

Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAATAGA 540

Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAATAGA 540

Qy 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAAAAACCCAGACGCC 600

Db 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAAAAACCCAGACGCC 600

Qy 601 CTGCTCTGCATATCTGAGAGGCTTCCAGAGCGGATCGTGGAAATGGTCTTTCGCAATCA 660

Db 601 CTGCTCTGCATATCTGAGAGGCTTCCAGAGCGGATCGTGGAAATGGTCTTTCGCAATCA 660

Qy 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCCT 720

Db 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCCT 720

Qy 721 CATGAATTTATTTGGGACGACATAAGGTGCTGTGCAGAAATGAACCTGGCGAGGAATGC 780

Db 721 CATGAATTTATTTGGGACGACATAAGGTGCTGTGCAGAAATGAACCTGGCGAGGAATGC 780

Qy 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACACATTTGACCAATTTAT 840

Db 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACACATTTGACCAATTTAT 840

Qy 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Qy 901 GGGCTCACCTGGGAATTAGAAACAAAGCACCTCGAGGGGCAACTACTTTGAGATGAGT 960

Db 901 GGGCTCACCTGGGAATTAGAAACAAAGCACCTCGAGGGGCAACTACTTTGAGATGAGT 960

Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTGCTGCTGCTGCTGCTGCT 1020

Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTGCTGCTGCTGCTGCTGCT 1020

Qy 1021 AGAAGCAGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

Db 1021 AGAAGCAGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

Qy 1081 GTTACCATCTGAGGAAAGGATTTATAATGCTACCAATTAAGTCAAGTCAAGATATGAAAT 1140

Db 1081 GTTACCATCTGAGGAAAGGATTTATAATGCTACCAATTAAGTCAAGTCAAGATATGAAAT 1140

Qy 1141 GACCAATATGAAGAGTGTGTTTCTGTGAGGCTTAAAGCCTACCCACAAATCAGATGT 1200

Db 1141 GACCAATATGAAGAGTGTGTTTCTGTGAGGCTTAAAGCCTACCCACAAATCAGATGT 1200

Qy 1201 ACGTGGACCTTCTCTCGAAATCATTTCCCTGTGAGCAAAAGGGTCTTGTATACGGATAC 1260

Db 1201 ACGTGGACCTTCTCTCGAAATCATTTCCCTGTGAGCAAAAGGGTCTTGTATACGGATAC 1260

Qy 1261 AGCATATCCAAGTGTGCAATCATTAAGCACAGCAGGAGAGATATATATTCATGAGAA 1320

Db 1261 AGCATATCCAAGTGTGCAATCATTAAGCACAGCAGGAGAGATATATATTCATGAGAA 1320

Qy 1321 AATGATGATGCCCAATTTACCAAAATGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

Db 1321 AATGATGATGCCCAATTTACCAAAATGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

Qy 1381 CTCGAGAGCATCGGCAAGTCCAGGGTCTGTTCTCGATGGATACCATTTACCATCT 1440

Db 1381 CTCGAGAGCATCGGCAAGTCCAGGGTCTGTTCTCGATGGATACCATTTACCATCT 1440

Qy 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATACAGAAAGGA 1500

Db 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATACAGAAAGGA 1500

Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGAGAGTGGGTGTCGAGCAGTACTCTA 1560

Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGAGAGTGGGTGTCGAGCAGTACTCTA 1560

Qy 1561 AACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGCTGATACAAATTCCTTGGC 1620

Db 1561 AACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGCTGATACAAATTCCTTGGC 1620

Qy 1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680

Db 1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680

Qy 1681 ATCTCATTTCTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

Db 1681 ATCTCATTTCTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

source

```
1. 3475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="13q12"
/cell_type="stem cell"
/tissue_type="bone marrow"
58. 3039
CDS
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="AAAI18947.1"
/db_xref="GI:409573"
/translation="MPALADRGATVPLLVFESAMIFGTTITNODLPVVKCVLINKNND
SSVKGSSYPWVSPEDLGCALRQSSGTVEEAAVEVDVSAISITQLVLVDAPGNIS
CLAVFKHSLNCPKPHDLQNGVSMVILKMTETQAGEYLLFPIOSEATNYILFTVSI
RNTLLYLRLPRYKFMNODALVCISYPEPIVENVLDCDSQGESKEESPVAVKKEE
KYHLHFGLDRCRCARNEDELRECFRTIDLPQFLKVEPGLWIRKAVH
VHFGGLTWELNKALKEGNYFEMSTYSTRMIRILFAFVSSVARNDTYTSSSK
HQSQALVTVISKGFINKHSPSEDEIDQYBEFCFSVRFKAYPOIRKCTWFSRKSFPK
EOKGLDNGYSISKFNHKGPEVIFHAENDDAOFTKMTLINIRKKPOVLAESAQA
SCFSDGYPLSPWTKKCDKSPNCTEELTEGVWNRKANRVFGGWVSSSTLNMSEAIAK
GFLVKCAVNSLGTSCETILLNSPGPFFIODNISFYATIGVCLLFIYVLLILCHKY
KKOFRYSQLOMVOVTOSSDNEYFYDFREYEDLKWEPFRENLEFKGLSGAFKV
MNATAYGISRTGYSIOAVKMLKEKADSSEREALMSLKMNTQLGSHENIVNLGACT
LSGPIYILFYCCYGDLLNLYLRKREKFRHTWTFIKFHNFSFYPTOSHNSMPGS
LGVQIHPSDQISGLHNSPHSDEIEYENOKRLEEREDLNVLTFEOLLCFAYOVAK
MERLEFKSVYHRDLAARNLVYTHGKVKIKCDFGLARDIMSDNSYVVRGNARLPVKWMA
PESLFEIGIYTIKSDWSYGLLLWEIFSLGVNPNYPGIPVDANFYKLIQNGPKMDQPFYA
TEEIIYIMQSCWAFDPSRKRSPFNLTSLFGLCOLADAEAMVQNDGRVSECPHTYQNR
RPSREMDLGLLSQAQVEDS"
BASE COUNT 1042 a 709 c 784 g 940 t
ORIGIN
```

Query Match 99.3%; Score 3475; DB 9; Length 3475;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGAGCGGCATCCGAGGGCTGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
DB	1	CGAGCGGCATCCGAGGGCTGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
QY	61	CCGCGGTTGGCGCGCAGCGGGACCGGTGCGCTGCTGTTGTTTCTGCAATGATA	120
DB	61	CCGCGGTTGGCGCGCAGCGGGACCGGTGCGCTGCTGTTGTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAAATCAAGATCTCGCTGTGATCAAGTGTTGTTTAAATCAATCAAG	180
DB	121	TTTGGGACTATTACAAATCAAGATCTCGCTGTGATCAAGTGTTGTTTAAATCAATCAAG	180
QY	181	AACATGATTCATCAGTGGGGAAGTCATCATATCCCATGATCAGATCCCGGAA	240
DB	181	AACATGATTCATCAGTGGGGAAGTCATCATATCCCATGATCAGATCCCGGAA	240
QY	241	GACCTCGGGTGTGCTTTCAGACCCCGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
DB	241	GACCTCGGGTGTGCTTTCAGACCCCGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
QY	301	GAAGTGGATGATCTGCTCCATCACACTGCAAGTGTGTCGATGCCCCAGGGAACATT	360
DB	301	GAAGTGGATGATCTGCTCCATCACACTGCAAGTGTGTCGATGCCCCAGGGAACATT	360
QY	361	TCCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAATTCAGGACACATTTTGAATTACAA	420
DB	361	TCCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAATTCAGGACACATTTTGAATTACAA	420
QY	421	AACAGAGGAGTTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
DB	421	AACAGAGGAGTTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
QY	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAGTATAGA	540
DB	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAGTATAGA	540
QY	541	AATACCTTGTCTTACACATTAAAGAGACCTTACTTTAGAAAAATGGAAACCCAGGACGC	600

DB	541	ATAACCTTGTCTTACACATTAAAGAGACCTTACTTTAGAAAAATGGAAACCCAGGACGC	600
QY	601	CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGCATGCTGGAATGGTGTCTTTCGCAATCA	660
DB	601	CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGCATGCTGGAATGGTGTCTTTCGCAATCA	660
QY	661	CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAGTCTT	720
DB	661	CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAGTCTT	720
QY	721	CATGAATTTATTTGGGACGACATAAGTGCTGTGCCAGAAATGAACCTGGCGGGAATGC	780
DB	721	CATGAATTTATTTGGGACGACATAAGTGCTGTGCCAGAAATGAACCTGGCGGGAATGC	780
QY	781	ACCGGCTGTTTACAAATAGATCTAAATCAAACCTCTCAGACACACATTGGCACAATTTT	840
DB	781	ACCGGCTGTTTACAAATAGATCTAAATCAAACCTCTCAGACACACATTGGCACAATTTT	840
QY	841	CTTAAGTAGGGGAACCTTATGGATTAAGTGCAAGCTGTTCATGTGAACCATGGATTC	900
DB	841	CTTAAGTAGGGGAACCTTATGGATTAAGTGCAAGCTGTTCATGTGAACCATGGATTC	900
QY	901	GGGCTCACCTCGGAATTAGAAACAAAGCACCTCGAGGAGGCAACTACTTTGAGATGAGT	960
DB	901	GGGCTCACCTCGGAATTAGAAACAAAGCACCTCGAGGAGGCAACTACTTTGAGATGAGT	960
QY	961	ACCTATTCAACAAACAGAACTATGATACGGAATCTTGTGCTTTTATCATCATCAGTGGCA	1020
DB	961	ACCTATTCAACAAACAGAACTATGATACGGAATCTTGTGCTTTTATCATCATCAGTGGCA	1020
QY	1021	AGAAACGACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTTG	1080
DB	1021	AGAAACGACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTTG	1080
QY	1081	GTTACCATCTGAGGAAGGATTTTAAATGCTACCAATTCAGTCAAGATATGATAAAT	1140
DB	1081	GTTACCATCTGAGGAAGGATTTTAAATGCTACCAATTCAGTCAAGATATGATAAAT	1140
QY	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGGTTTAAAGCCTTACCACAAATCAGATGT	1200
DB	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGGTTTAAAGCCTTACCACAAATCAGATGT	1200
QY	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGAGCAAAAGGCTTGTGATAACGGATAC	1260
DB	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGAGCAAAAGGCTTGTGATAACGGATAC	1260
QY	1261	AGCATATCCAAGTTTTCGAATCATAGCACCCAGGAGATATATATTTCCATGCGAGAA	1320
DB	1261	AGCATATCCAAGTTTTCGAATCATAGCACCCAGGAGATATATATTTCCATGCGAGAA	1320
QY	1321	AATGATGATGCCAATTTTACCATAATGTTTCAGCTGAATATAGAGAGAAACCTCAAGTG	1380
DB	1321	AATGATGATGCCAATTTTACCATAATGTTTCAGCTGAATATAGAGAGAAACCTCAAGTG	1380
QY	1381	CTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTTACCATCT	1440
DB	1381	CTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTTACCATCT	1440
QY	1441	TGGACCTCGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATACAGAGGA	1500
DB	1441	TGGACCTCGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATACAGAGGA	1500
QY	1501	GTCGTGAATAGAAAGGCTAAACAGAAAAGTGTGACAGTGGTGTGCGAGCAGTACTCTA	1560
DB	1501	GTCGTGAATAGAAAGGCTAAACAGAAAAGTGTGACAGTGGTGTGCGAGCAGTACTCTA	1560
QY	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAAATTCCTTGGC	1620
DB	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAAATTCCTTGGC	1620
QY	1621	ACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCAAGACAAC	1680
DB	1621	ACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCAAGACAAC	1680

Db 1621 ACATCTTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680
Qy 1681 ATCTCATCTTATGCAACAATTTGGTGTGCTCTCTCTTCATTTGCTGTTTAAACCTCGTA 1740
Db 1681 ATCTCATCTTATGCAACAATTTGGTGTGCTCTCTCTTCATTTGCTGTTTAAACCTCGTA 1740
Qy 1741 ATTTGTCTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTCTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCTCTCAGATAATAGTACTTCTACGTTTGGGAAGGTACTAGGATCATGATATGATCTC 1860
Db 1801 ACCGGCTCTCTCAGATAATAGTACTTCTACGTTTGGGAAGGTACTAGGATCATGATATGATCTC 1860
Qy 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCATGATATGATCTC 1920
Db 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCATGATATGATCTC 1920
Qy 1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
Db 1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
Qy 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Qy 2041 CTCAAGATGATGCCAGCTGGGAAGCCACAGAGATATTGTGAACCTGCTGGGGCGTGC 2100
Db 2041 CTCAAGATGATGCCAGCTGGGAAGCCACAGAGATATTGTGAACCTGCTGGGGCGTGC 2100
Qy 2101 ACAGTGTCTCAGACCAATTTACTTGAATTTTGAATCTGCTGATGTTGATCTTCTCAAC 2160
Db 2101 ACAGTGTCTCAGACCAATTTACTTGAATTTTGAATCTGCTGATGTTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGAGCTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGAGCTTGGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTCAAGTTTACCCCACTTTTCAATTCATCAATCAATCAATCAATCAATCAATCAATCAAT 2280
Db 2221 AATTTCAAGTTTACCCCACTTTTCAATTCATCAATCAATCAATCAATCAATCAATCAATCAAT 2280
Qy 2281 GAAGTTCAAGTACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAAC 2340
Db 2281 GAAGTTCAAGTACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTCTTTGCTTGTGATATCAAGTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTCTTTGCTTGTGATATCAAGTTGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGAGGAGCTGCTGTCACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGAGGAGCTGCTGTCACCCAC 2520
Qy 2521 GGAAGAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTTCCAAAC 2580
Db 2521 GGAAGAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTTCCAAAC 2580
Qy 2581 TATGTTGTACAGGGAATGCCCGTGTGCTGTAAATGATGGCCCCGAAAGCCCTGTTT 2640
Db 2581 TATGTTGTACAGGGAATGCCCGTGTGCTGTAAATGATGGCCCCGAAAGCCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCAATTAAGAGTATGTTGCTCATATGGAATATTACTTGTGGGAATTC 2700
Db 2641 GAAGGCATCTACACCAATTAAGAGTATGTTGCTCATATGGAATATTACTTGTGGGAATTC 2700
Qy 2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTTGATGCTAACTTCTACAACATG 2760
Db 2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTTGATGCTAACTTCTACAACATG 2760

Qy 2761 ATTCAAAATGGATTTAAAAATGATCAGCATTTTATGCTACAGAGAATATACATTATA 2820
Db 2761 ATTCAAAATGGATTTAAAAATGATCAGCATTTTATGCTACAGAGAATATACATTATA 2820
Qy 2821 ATGCAATCTCTCTGGGCTTTTGAATCAAGAAAGCCGATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTCTCTGGGCTTTTGAATCAAGAAAGCCGATCTTCCCTAAATTTGACTTCG 2880
Qy 2881 TTTTGTAGATCTCAGCTGGCAGATGCCAAGAAAGGATGTATCAGAAATGTGATGCCGT 2940
Db 2881 TTTTGTAGATCTCAGCTGGCAGATGCCAAGAAAGGATGTATCAGAAATGTGATGCCGT 2940
Qy 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCCAGCAGAGATGGATTG 3000
Db 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCCAGCAGAGATGGATTG 3000
Qy 3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAAGTTCGTAGAGAAATTTAGTTTAAAG 3060
Db 3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAAGTTCGTAGAGAAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Db 3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Qy 3121 CACTAAAAGAAAATCTATTATCACTGCTTCCAGCAGCTTTTCTCTAGAGCGCTCT 3180
Db 3121 CACTAAAAGAAAATCTATTATCACTGCTTCCAGCAGCTTTTCTCTAGAGCGCTCT 3180
Qy 3181 CGGTTTACTCTTTTCAAGGAGCTTTTGTAAATCAAAATCATCTCTCACAAGGCAG 3240
Db 3181 CGGTTTACTCTTTTCAAGGAGCTTTTGTAAATCAAAATCATCTCTCACAAGGCAG 3240
Qy 3241 GAGGAGCTGATTAATGAACCTTTTATGGAGCATTTGATCGATCCAGGCGCTTCTCAGGCG 3300
Db 3241 GAGGAGCTGATTAATGAACCTTTTATGGAGCATTTGATCGATCCAGGCGCTTCTCAGGCG 3300
Qy 3301 GCTTCAGTGAATTTGTACCTCAAGTACAGTATTTCTTGTAAATCAATAAACAAGC 3360
Db 3301 GCTTCAGTGAATTTGTACCTCAAGTACAGTATTTCTTGTAAATCAATAAACAAGC 3360
Qy 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
Db 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
Qy 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475
Db 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475

RESULT 7

140109 LOCUS 140109 3476 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 1 from patent US 5618709.
ACCESSION I40109
VERSION I40109.1 GI:2083114
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3476)
AUTHORS Gewirtz, A.M., Small, D. and Civin, C.I.
TITLE Antisense oligonucleotides specific for SPK-1 and method for
inhibiting expression of the SPK-1 protein
JOURNAL Patent: US 5618709-A 1 08-APR-1997;
FEATURES Location/Qualifiers
Source 1. .3476
BASE COUNT 1044 a 709 c 783 g 940 t
ORIGIN
Query Match 99.2%; Score 3474.4; DB 6; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3475: Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	1	CGAGCGGCATCCGAGGCTGGCGCGGCTGGGGAGACCCGGCTCCGGAGCCATG	60	
Db	1	CGAGCGGCATCCGAGGCTGGCGCGGCTGGGGAGACCCGGCTCCGGAGCCATG	60	
Qy	61	CCGCGTTGGCGCGGACCGGCGGACCGTGCCTGCTGTTTTCGCAATGATA	120	
Db	61	CCGCGTTGGCGCGGACCGGCGGACCGTGCCTGCTGTTTTCGCAATGATA	120	
Qy	121	TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG	180	
Db	121	TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG	180	
Qy	181	AACAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA	240	
Db	181	AACAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA	240	
Qy	241	GACCTCGGGTGTGGTTGAGACCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300	
Db	241	GACCTCGGGTGTGGTTGAGACCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300	
Qy	301	GAAGTGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGAAACATT	360	
Db	301	GAAGTGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGAAACATT	360	
Qy	361	TCCTGTCTCGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420	
Db	361	TCCTGTCTCGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420	
Qy	421	AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCAAGCTGGAGAATAC	480	
Db	421	AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCAAGCTGGAGAATAC	480	
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTACAGTCAAGTATAAGA	540	
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTACAGTCAAGTATAAGA	540	
Qy	541	AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCGCC	600	
Db	541	AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCGCC	600	
Qy	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGTGTTCGCAATCA	660	
Db	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGTGTTCGCAATCA	660	
Qy	661	CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGCTT	720	
Db	661	CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGCTT	720	
Qy	721	CATGAATTTTGGGACGGACATGAAGTGTGTCGACAGAAATGAACCTGGGAGGAATGC	780	
Db	721	CATGAATTTTGGGACGGACATGAAGTGTGTCGACAGAAATGAACCTGGGAGGAATGC	780	
Qy	781	ACCAGGCTGTCACAATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAATTTATTT	840	
Db	781	ACCAGGCTGTCACAATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAATTTATTT	840	
Qy	841	CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAAAGCTGTTCATGTGAACCATGGATTC	900	
Db	841	CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAAAGCTGTTCATGTGAACCATGGATTC	900	
Qy	901	GGGCTACCTGGGAAATTAGAAACAAAGCACTCAGGAGGGCACTACTTTGAGATGAGT	960	
Db	901	GGGCTACCTGGGAAATTAGAAACAAAGCACTCAGGAGGGCACTACTTTGAGATGAGT	960	
Qy	961	ACCTATTCAACACACAGACTATCATACGGATTCTGTTGCTTTGTATCATCATCACTGCA	1020	
Db	961	ACCTATTCAACACACAGACTATCATACGGATTCTGTTGCTTTGTATCATCATCACTGCA	1020	
Qy	1021	AGAAACGACACCGGATACTACACTTTGTCCTCTCAAGAGATCCAGTCAATCAGCTTTG	1080	
Db	1021	AGAAACGACACCGGATACTACACTTTGTCCTCTCAAGAGATCCAGTCAATCAGCTTTG	1080	

Qy	1081	GTTACCATCGTAGGAAAGGATTTTATAATGCTAGCAATTTCAAGTGAAGATTATGAAATT	1140	
Db	1081	GTTACCATCGTAGGAAAGGATTTTATAATGCTAGCAATTTCAAGTGAAGATTATGAAATT	1140	
Qy	1141	GACCAATATGAAGAGTTTTTGTCTGTCTAGGTTTAAAGCCTACCCACAAATCAGATGT	1200	
Db	1141	GACCAATATGAAGAGTTTTTGTCTGTCTAGGTTTAAAGCCTACCCACAAATCAGATGT	1200	
Qy	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGAGCAAAAGGGTCTTGTATACGGATAC	1260	
Db	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGTGACAAAAGGGTCTTGTATACGGATAC	1260	
Qy	1261	AGCATATCCAAGTTTTTGAATCATTAAGCACACGAGGAGAAATATATATTTCCATGCAGAA	1320	
Db	1261	AGCATATCCAAGTTTTTGAATCATTAAGCACACGAGGAGAAATATATATTTCCATGCAGAA	1320	
Qy	1321	AATGATGATGCCAATTTTACCAAAATGTTACGCTGAATATAGAAGAGAAACCTCAAGTG	1380	
Db	1321	AATGATGATGCCAATTTTACCAAAATGTTACGCTGAATATAGAAGAGAAACCTCAAGTG	1380	
Qy	1381	CTCGAGAGCATCGGCAAGTCAGGCTCCTGTTCTCGGATGGATACCCATTTACCATCT	1440	
Db	1381	CTCGAGAGCATCGGCAAGTCAGGCTCCTGTTCTCGGATGGATACCCATTTACCATCT	1440	
Qy	1441	TGGACCTCGAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAAGAGATCACAGAAGGA	1500	
Db	1441	TGGACCTCGAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAAGAGATCACAGAAGGA	1500	
Qy	1501	GTCCTGGAATAGAAAGGCTTAACAGAAAAGTGTGTCGACAGTGGGTGTCGACAGTACTCTA	1560	
Db	1501	GTCCTGGAATAGAAAGGCTTAACAGAAAAGTGTGTCGACAGTGGGTGTCGACAGTACTCTA	1560	
Qy	1561	AACATGAGTGAAGCCATAAAAGGTTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC	1620	
Db	1561	AACATGAGTGAAGCCATAAAAGGTTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC	1620	
Qy	1621	ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTTATCAAGACAAC	1680	
Db	1621	ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTTATCAAGACAAC	1680	
Qy	1681	ATCTCATTTCTATGCAACAAATTTGTTCTCTCTTTCATTTCTGCTTTTAAACCTTGCTA	1740	
Db	1681	ATCTCATTTCTATGCAACAAATTTGTTCTCTCTTTCATTTCTGCTTTTAAACCTTGCTA	1740	
Qy	1741	ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTACAGGTG	1800	
Db	1741	ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTACAGGTG	1800	
Qy	1801	ACCGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATCAATATGATCTC	1860	
Db	1801	ACCGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATCAATATGATCTC	1860	
Qy	1861	AAATGGGAGTTTCCAAAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920	
Db	1861	AAATGGGAGTTTCCAAAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920	
Qy	1921	GGAAAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTT	1980	
Db	1921	GGAAAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTT	1980	
Qy	1981	GCCGTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040	
Db	1981	GCCGTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040	
Qy	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTTGAACTGCTGGGGGCTGC	2100	
Db	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTTGAACTGCTGGGGGCTGC	2100	
Qy	2101	ACACTGTCAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC	2160	
Db	2101	ACACTGTCAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC	2160	

Qy	958	AGTACCTATTCAACAACAGAACTATGATACGAGTCTCTGTTTGGTTTGTATCATCATCAGT	101
Db	901	AGTACCTATTCAACAACAGAACTATGATACGAGTCTCTGTTTGGTTTGTATCATCATCAGT	960
Qy	1018	GCAAGAAACGACACCGGATACTACACTTGTTTCCTCTTCAAGCATCCAGTCAATCAGCT	1077
Db	961	GCAAGAAACGACACCGGATACTACACTTGTTTCCTCTTCAAGCATCCAGTCAATCAGCT	1020
Qy	1078	TTGGTTTACCATCGTAGGAAAGGGATTATAAATGCTACCAATTCAAGTGAAGATTATGAA	1137
Db	1021	TTGGTTTACCATCGTAGGAAAGGGATTATAAATGCTACCAATTCAAGTGAAGATTATGAA	1080
Qy	1138	ATTGACCAATATGAAGAGTTTGTGTTTCTGTGTCAGGTTTAAAGCCTACCACAAATCAGA	1197
Db	1081	ATTGACCAATATGAAGAGTTTGTGTTTCTGTGTCAGGTTTAAAGCCTACCACAAATCAGA	1140
Qy	1198	TGTACGTGGACCTTCTCTCGAAATCATTTCCCTGTGCAGCAAAAGGCTTTGATAACGGA	1257
Db	1141	TGTACGTGGACCTTCTCTCGAAATCATTTCCCTGTGCAGCAAAAGGCTTTGATAACGGA	1200
Qy	1258	TACAGCATATCCAAAGTTTTCGAATCATAAAGCACAGCCAGGAGGAATATATATTCATGCA	1317
Db	1201	TACAGCATATCCAAAGTTTTCGAATCATAAAGCACAGCCAGGAGGAATATATATTCATGCA	1260
Qy	1318	GAAATATGATGATGCCAATTTTACCAAAATGTTTCACGCTGAATATAAGAAGAAACCTCAA	1377
Db	1261	GAAATATGATGATGCCAATTTTACCAAAATGTTTCACGCTGAATATAAGAAGAAACCTCAA	1320
Qy	1378	GTGCTCCGCAAGCATCGGAAGTCTCAGACAAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1437
Db	1321	GTGCTCCGCAAGCATCGGAAGTCTCAGACAAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1380
Qy	1438	TCATTGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1497
Db	1381	TCATTGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1440
Qy	1498	GGAGTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGGACAGTGGGTGTCGAGCAGTACT	1557
Db	1441	GGAGTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGGACAGTGGGTGTCGAGCAGTACT	1500
Qy	1558	CTAAACATGAGTGAAGCCATAAAGGGTTCTTGTCGAAGTCTGTGTCATACAAATTCCTCTT	1617
Db	1501	CTAAACATGAGTGAAGCCATAAAGGGTTCTTGTCGAAGTCTGTGTCATACAAATTCCTCTT	1560
Qy	1618	GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC	1677
Db	1561	GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC	1620
Qy	1678	AACATCTCATTCATGCAACAAATGGTGTGTTGTCCTCTTTCATATGTCGTTTTAACCTGT	1737
Db	1621	AACATCTCATTCATGCAACAAATGGTGTGTTGTCCTCTCTCTCTCTCTCTCTCTCTCTCT	1680
Qy	1738	CTAATTTGTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
Db	1681	CTAATTTGTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
Qy	1798	GTGACCGGCTCCTCAGATAATGAGTACTCTTACCTGTTGATTTTCAGAGAAATATGAATATGAT	1857
Db	1741	GTGACCGGCTCCTCAGATAATGAGTACTCTTACCTGTTGATTTTCAGAGAAATATGAATATGAT	1800
Qy	1858	CTCAAAATGGGAGTTTCCCAAGAGAAATTTAGAGTTTGGGAGGCTACTAGGATCAGGTGCT	1917
Db	1801	CTCAAAATGGGAGTTTCCCAAGAGAAATTTAGAGTTTGGGAGGCTACTAGGATCAGGTGCT	1860
Qy	1918	TTTGGAAAAGTGATGAACGCAACAGCTTATGGAAATTAGCAAAACAGGAGTCTCAATTCAG	1977
Db	1861	TTTGGAAAAGTGATGAACGCAACAGCTTATGGAAATTAGCAAAACAGGAGTCTCAATTCAG	1920
Qy	1978	GTTGCCCTCAAAATGCTCAAGAAAGAACAGACAGCTCTGAAAGAGGCACTCATGTCA	2037
Db	1921	GTTGCCCTCAAAATGCTGAAAGAAAGAACAGACAGCTCTGAAAGAGGCACTCATGTCA	1980
Qy	2038	GAACCTAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTGTGAACTCTGCTGGGGGCG	2097

361	DB	CAAAACAGAGGAGTGTGTTTCCATGGTGCATTTTGAANAATGACAGAAACCCCAAGCTGGAGAA	420
478	QY	TACCTACTCTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTCAAGTATA	537
421	DB	TACCTACTCTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTCAAGTATA	480
538	QY	AGAAATACCTCGCTTTACACATTTAAGNAGACCTTACTTTTGAANAATGGAANAACCAAGGAC	597
481	DB	AGAAATACCTCGCTTTACACATTTAAGNAGACCTTACTTTTGAANAATGGAANAACCAAGGAC	540
598	QY	GCCTGTGCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGCGAT	657
541	DB	GCCTGTGCTGCATATCTGAGAGCGTTCCAGAGCGGATCTCGAAATGGGTGCTTTTGGCGAT	600
658	QY	TCACAGGGGGAAAGCTGTAAAGAAAGATCCAGCTGTTGTTTAAAAGGAGGAAAAAGTG	717
601	DB	TCACAGGGGGAAAGCTGTAAAGAAAGATCCAGCTGTTGTTTAAAAGGAGGAAAAAGTG	660
718	QY	CTTCATGAATTTATTTGGGACGGACATTAAGTGTCTGCGCAGAAATGAACHTGGCGAGGGAA	777
661	DB	CTTCATGAATTTATTTGGGACGGACATTAAGTGTCTGCGCAGAAATGAACHTGGCGAGGGAA	720
778	QY	TGCACACAGGCTGTTCACAATTAGATCTTAATCAAACTCCTCAGACCACTTGCACCAATTA	837
721	DB	TGCACACAGGCTGTTCACAATTAGATCTTAATCAAACTCCTCAGACCACTTGCACCAATTA	780
838	QY	TTTTCTTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGA	897
781	DB	TTTTCTTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGA	840
898	QY	TTGCGGCTCACCTGGGAAATTTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTTGAGATG	957
841	DB	TTGCGGCTCACCTGGGAAATTTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTTGAGATG	900
958	QY	AGTACCTTATTCAAACACAGCACTATCATACGGATTCTGTTTCTGTTTGTATCATCATG	1017
901	DB	AGTACCTTATTCAAACACAGCACTATCATACGGATTCTGTTTCTGTTTGTATCATCATG	960
1018	QY	GCAAGAAACGACACCGGATACTACACTTGTTCCTCTTTCAAAGCATCCAGTCAATCAGCT	1077
961	DB	GCAAGAAACGACACCGGATACTACACTTGTTCCTCTTTCAAAGCATCCAGTCAATCAGCT	1020
1078	QY	TTGGTTTACCATTCTAGGAAAGGATTTATAAATGCTTACCAATTTCAAGTGAAGATTTATGAA	1137
1021	DB	TTGGTTTACCATTCTAGGAAAGGATTTATAAATGCTTACCAATTTCAAGTGAAGATTTATGAA	1080
1138	QY	ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAAATCAGA	1197
1081	DB	ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAAATCAGA	1140
1198	QY	TGTACGTGGACCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTGTATAACGGA	1257
1141	DB	TGTACGTGGACCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTGTATAACGGA	1200
1258	QY	TACAGCATATCCAAAGTTTTCGAATCATAAAGCACCGCAGGAGGAATATATTTCCATGCA	1317
1201	DB	TACAGCATATCCAAAGTTTTCGAATCATAAAGCACCGCAGGAGGAATATATATTTCCATGCA	1260
1318	QY	GAATAATCATATGCCCAATTTACCAAAATGTTACGCTGAATATAAGAAGGAACCTCAA	1377
1261	DB	GAATAATCATATGCCCAATTTACCAAAATGTTACGCTGAATATAAGAAGGAACCTCAA	1320
1378	QY	GTGCTCGCAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTACCA	1437
1321	DB	GTGCTCGCAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTACCA	1380
1438	QY	TCTTTGGACCTGGAGAAAGTGTTCAGACAAGTCTCTCCAACTGCACAGAAGAGATCACAGAA	1497
1381	DB	TCTTTGGACCTGGAGAAAGTGTTCAGACAAGTCTCTCCAACTGCACAGAAGAGATCACAGAA	1440
1498	QY	GGAGTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTTGCGACAGTGGGTGTCGACGAGTACT	1557
1441	DB	GGAGTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTTGCGACAGTGGGTGTCGACGAGTACT	1500

QY 2638 TTTGAAGCATCTACACCATTAAGAGTGATGTCTGGTCATATGAATATTACTGTGGAA 2697
Db 2581 TTTGAAGCATCTACACCATTAAGAGTGATGTCTGGTCATATGAATATTACTGTGGAA 2640
QY 2698 ATCTTCTCAGTGTGTGAATCCTTACCCTGGCAATCCGGTGTAGTGTAATCTTCAAAA 2757
Db 2641 ATCTTCTCAGTGTGTGAATCCTTACCCTGGCAATCCGGTGTAGTGTAATCTTCAAAA 2700
QY 2758 CTGATTCAAAATGATTTAAATGGATCAGCCATTTTATGCTACAGAGAATATACATT 2817
Db 2701 CTGATTCAAAATGATTTAAATGGATCAGCCATTTTATGCTACAGAGAATATACATT 2760
QY 2818 ATAAATGAATCCTCTGGGCTTTTGAATCAAGGAAGGCCATCTTCCCTAAATTTGACT 2877
Db 2761 ATAAATGAATCCTCTGGGCTTTTGAATCAAGGAAGGCCATCTTCCCTAAATTTGACT 2820
QY 2878 TCGTTTTTAGGATGTCTAGCTGGCAGATCGAAGAAGCGATGTATCAGAATGTGGATGGC 2937
Db 2821 TCGTTTTTAGGATGTCTAGCTGGCAGATCGAAGAAGCGATGTATCAGAATGTGGATGGC 2880
QY 2938 CGTGTTCGGAATGTCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2997
Db 2881 CGTGTTCGGAATGTCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2940
QY 2998 TTGGGGCTACTCTCCGAGGCTCAGCTCGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3057
Db 2941 TTGGGGCTACTCTCTCCGAGGCTCAGCTCGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3000
QY 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTT 3117
Db 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGTAAATTT 3060
QY 3118 CATCACTAAAAGAAATCTATATCACTGCTGTCTTACCAGACTTTTCTCTAGAGCGG 3177
Db 3061 CATCACTAAAAGAAATCTATATCACTGCTGTCTTACCAGACTTTTCTCTAGAGCGG 3120

RESULT 10
144732
LOCUS I44732 3521 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5635388.
ACCESSION I44732
VERSION I44732.1 GI:2469445
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3521)
AUTHORS Bennett, B.D., Broz, S.D., Matthews, W. and Zeigler, F.C.
TITLE Agonist antibodies against the flk2/flt3 receptor and uses thereof
JOURNAL Patent: US 5635388-A 1 03-JUN-1997;
FEATURES
Location/Qualifiers
1..3521
/organism="unknown"
BASE COUNT 943 a 854 c 907 g 817 t
ORIGIN

Query Match 65.0%; Score 2274.6; DB 6; Length 3521;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 2805; Conservative 0; Mismatches 649; Indels 41; Gaps 10;

QY 17 GCGTGGCGCGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATCGCGGGTGGC---GC 73
Db 42 GCGTGGCGCGCGCGCT-GGCTACCGCGCGCTCCGGAGGCCATCGCGGGCTGGCGCAG 100
QY 74 GCGAGCGGCGACCGGTGCGCGCTGCTGCTGTTTTTCTGCAATGATATTGGGACTATTA 133
Db 101 GCAGCGACCGCGCGGCTGCTGCTGTTGTTTGTCAAGTAATGATCTTTGAGACCGTTA 160
QY 134 CAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAGAAACAATGATTCAT 193
Db 161 CAAACCAAGACCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATGATGAGCAATGGCTCAT 220

QY 194 CAGTGGGAAGTCAATCATATATCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTG 253
Db 221 CAGCGGGAAGCCATCATCTGTAACGAATGGTGGAGGATCCCGAAGACCTCCAGTGA 280
QY 254 CGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCGGCTGTGGAAGTGGATGAT 313
Db 281 CCCGAGCGCGCAGAGTGAAGGACGGTATATGAAGCGCGCCACCCTGGAGGTGGCCGAGT 340
QY 314 CTGCTTCATCACACTGCAAGTGTGTGCTGATCCCGAGGAACAATTTCTCTCTCTGG 373
Db 341 CTGGGTCAATCACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTCTCTGG 400
QY 374 TCTTTAAGCACAGCTCCCTGAATTTGCCAGCCACATTTTGAATTTACAAACAGAGAGTGT 433
Db 401 TCTTTAAGCACAGCTCCCTGGGCTGCCAGCGCACTTTTGAATTTACAAACAGAGAAATCG 460
QY 434 TTTCCATGGTCATTTTGAATTTGAAATGACAGAAACCCAGCTGGAGAATACCTTTTATTC 493
Db 461 TTTCCATGGCTCTTTGAACGTGACAGAGACCCAGGAGGAATACCTTACCTCATTTTC 520
QY 494 AGAGTGAAGCTACCAATTAACAATTTGTTTACAGTGAATATAAGAAATACCTCTCTTT 553
Db 521 AGAGCGAAGCGCCCACTACACAGTACTGTTTCCACAGTGAATGTAGAGATACACACTGT 580
QY 554 ACACATTAAGAGACCTTACTTTAGAAAATGGAACACAGACGCGCTCTGTCATAT 613
Db 581 ATGCTCTAAGGAGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCACT 640
QY 614 CTGAGAGCGTTCCAGAGCGGATCGTGGATGGTGTCTTCCGATTTCCAGAGGGGAAAGCT 673
Db 641 CGAGGGGTGTTCCGGAGCGCCACTGTGGAGTGGTGTCTGCACTCCCAAGGAAAGCT 700
QY 674 GTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGTCTATGAATTTTG 733
Db 701 GTAAAGAGAAAGCGCTGTGTGTCAGAAAGAGGAGAAAGTGTCTATGAGTGTTCG 760
QY 734 GGAGGACATAAGTGTCTGCCAGAAATGAATGGGAGGGAATGACACAGGCTGTTCAT 793
Db 761 GAACAGACATCAGATGCTGTGTAGAAATGCACTGGCGCGGAATGACCAAGCTGTTC 820
QY 794 CAATAGATTAATCAAACTCTCAGACCACTGCGCAACATTTATTTCTTAAAGTAGGG 853
Db 821 CCATAGATCAAAACAGGCTCTCAGACCACTGCGCCAGTATTCTTCCGAAAGTGGGG 880
QY 854 AACCTTATGGATAGGTGCAAGCTGTTTCAATGAACCATGGATTCGGCTCACCTGGG 913
Db 881 AACCTTGTGGATCAGGTGTAGGCCATCCATGTGAACCATGGATTCGGCTCACCTGGG 940
QY 914 AATTAGAAAACAAGACCTCAGGAGGCGCACTTTTGAGATGAGTACCTATTCAACAA 973
Db 941 AGCTGGAAGACAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTACTCCCAA 1000
QY 974 ACAGAACTATGATACGGATTCGTTTGTATCATCAGTGGCAAGAAACGACACCG 1033
Db 1001 ACAGGACCATGATTCGGATTCCTTGGCTTTGTGCTTCCCGTGGGAAGAACGACACCG 1060
QY 1034 GATACATACCTTGTCTTCAAAAGCATCCCAAGTCAATCAGCTTTGGTTTACCATCTAG 1093
Db 1061 GATATTAACCTCTCTTCTCAAGCACCCAGCAGCTCAGCGTGTGGTGCACCATCTAG 1120
QY 1094 GAAAGGATTTAATAATGCTTACCAATTCAGTGAAGATTAATAATGACCAATATGAAG 1153
Db 1121 ARAAGGGTTTAAACGCTTACCAGCTCGCAAGAAGAGATGAATTTGACCCGTAACAAA 1180
QY 1154 AGTTTGTCTGTGTCAGGTTTAAAGCTTACCACCAAAATCAGATGACGTGGACCTTCT 1213
Db 1181 AGTTCTGCTTCTCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGACCGTGGATCTCT 1240
QY 1214 CTCGAAAATCATTTCTCTGTGACAAAGGCTCTTGATAACGGATACAGATATTCAGT 1273
Db 1241 CTCAGGCTCATTTCTCTGTGAAACAGAGGCTGGAGGATGGGTACAGCATATCTAAT 1300
QY 1274 TTTGCAATCATAAAGCACCGACCGAGGAGATATATATTTCCATGCAAGAAATGATGATGCC 1333


```

Qy 3471 TACAGAAAAA 3485
Db 3507 TGTAGAAAAA 3521

RESULT 11
LOCUS MMFLT3 3521 bp mRNA linear ROD 04-DEC-1991
DEFINITION Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF.
ACCESSION X59398
VERSION X59398.1 GI:50978
KEYWORDS Flt3 gene; tyrosine kinase receptor.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1321)
Rosnet, O.
Direct Submission
Submitted (07-MAY-1991) O. Rosnet, Cancerologie et therapeutique
exp., Inst National de la Sante, 27 Boulevard Lei Roure, 13009
Marseille, France
2 (bases 1 to 3521)
Rosnet, O., Marchetto, S., deLapeyriere, O. and Birnbaum, D.
Murine Flt3, a gene encoding a novel tyrosine kinase receptor of
the PDGFR/CSF1R family
Oncogene 6 (9), 1641-1650 (1991)
JOURNAL MEDLINE 92019834
PUBMED 1656368

FEATURES
Location/Qualifiers
1..3521
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3521
/gene="flt3"
1..3521
/gene="Flt3"
/evidence="experimental"
82..3084
/gene="Flt3"
/codon_start=1
/product="Flt3 protein"
/protein_id="CAA42041.1"
/db_xref="GI:50979"
/db_xref="MGD:MGI:95559"
/db_xref="SWISS-PROT:Q00342"
/translation="MRLAORSDRLLLLVLSVMILETVT NODLPVTKCVLISHENN
GSSAGKPSYRMVRS PEDLOCTPROSEGVYEATVEVAESGSITLQVOLATPGDL
SCLWVFKSSLCQPHFDLONRGIVSMALINVTQAGEYLLHIQSEANVTVLTVN
VRDTQLYLRRPFRKMNQDALLCISEGVEPTVEWVLCSSHRESKEEGPAVVRKE
EKVLHELFGRDRCARNALRGSTKLFITDLNAPQSTLPOLFLKVGDEPLWIRKAI
HVNHGFLTWELDEKLEGSYFEMSTYSTRMIRIILAFVSSVGRNDTGYTSSSS
KHPQSALVTILEKGFINATSSOEEVIDPYEKFCSVRFRAYPRICRTWIFSOAFSP
CEORLEGGYSTSKFCDHNKPGEVIFVAENDDAOFTKMTLINIRKPOVLANASASQ
ACSSDGPPLPSRWKKSXSDSPNCTEELPEIPEGVWKNKANKRVFGOWSSVTLNMEAG
KGLLVKXSNMTSGTSCETIFINSPGPPFIODNISFYATIGLCPLFIIVLVLICHK
YKQPFYSOLOMIQVTPLDNEYFYVDPRDYEDLKWPEFRENLEFGLVSGSAPGR
VNNATAYGISTKGTGVSQVAKMLKEKADSCERKALMSKLMTHLGHNDHNVNLLGAC
TSLGVPLIFCYCCGDLLNLRKREKFRHTWTEIFKEHNE SFPTFOAHNSNMPG
SREYVOLPDLQSLGNGNSIHSEDEIEYENOKRLAEEEDLNLVTFEDLLCFAYQV
AKGMFLEKSCVRDLARNVLVTHGKVVKICDFGLARDILSDSYVYVRGNARLPVK
WMAPELSLEGIYTIKSDVMSYGIILLWEIFSLGVNPPYGPVDPANFYKTLQSGVMEQP
FYATEGIYFVMSWAFDSKRSPFNLTSLFGLQALAEAEAEAMQNMGNGVPEHPSIY
QNRRLSREAGSEPPSPQAQVKIHRERS"
BASE COUNT 943 a 854 c 907 g 817 t
ORIGIN
Query Match 65.0%; Score 2274.6; DB 10; Length 3521;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 2805; Conservative 0; Mismatches 649; Indels 41; Gaps 10;
Qy 17 GCGTGGCGCGCCCTCGGGGACCCCGGCTCCGGAGGCCATGCCGGCGTTGGC 73

```

```

Db 42 GGCTGGCGCGCGCCT -GGCTACCGCGCTCCGGAGGCCATCGCGCGTGGCGCAGC 100
Qy 74 GCGACGGGGGACCGTGCCTGCTGCTGTTTCTGCAATGATATTGGGACTATTA 133
Db 101 GCAGCGCGCGCGCGCTGCTGCTGTTGTTTGTCTAGTAATGATCTTGAGACCGTTA 160
Qy 134 CAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATGAACAAATGATTCAT 193
Db 161 CAACCAAGACCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATGAACAAATGATTCAT 220
Qy 194 CAGTGGGGAAGTCAATCATATATCCATGATGATGATGATGATGATGATGATGATGATGATG 253
Db 221 CAGCGGGAAGCCATCATCGTACCGAATGGTGGAGGATCCCGGAGACCTCCAGTGTGA 280
Qy 254 CGTTGAGACCCAGAGCTCAGGAGCAGTGTAGGAAGCTGCCGCTGTGGAGGTGGATGAT 313
Db 281 CCCGAGCGCGCAGAGTGAAGGCGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGT 340
Qy 314 CTGCTTCATCACACTGCAAGTGTGTCGATGCCCGCAGGGAACATTTCTCTCTCTGGG 373
Db 341 CTGGGTCCATCACCTGCAAGTGCAGCTGCCACCCCGGAGGACCTTTCTCTCTCTGGG 400
Qy 374 TCTTTAAGCAGAGCTCCCTGGAATGCCAGCCACATTTTGAATTTACAAAACAGAGGATTCG 433
Db 401 TCTTTAAGCAGAGCTCCCTGCGCTGCCAGCCACCTTTGATTTACAAAACAGAGGAAATCG 460
Qy 434 TTTCCATGGTCATTTTGAATAATGACAGAAACCCAGCTGGAGAATACCTTCTTTTATTC 493
Db 461 TTTCCATGGCCATCTTTGAACGTGACAGAGACCCAGCAGGAGAAATACCTTCTCCATATTC 520
Qy 494 AGAGTGAAGCTACCAATACAAATATTTTACAGTGAATATTAAGAAATACCTCTCTTT 553
Db 521 AGAGCGAAGCCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGT 580
Qy 554 ACACATTAAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCGCCCTGCTGTGATAT 613
Db 581 ATGTGCTAAGGAGACCTTACTTTAGGAAGATGAAAAACAGGATGCTCTGCTGATCT 640
Qy 614 CTGAGAGCGTTCAGAGCGCGATCGTGAATGGTGTCTTGGGATTCACAGGGGGAAGCT 673
Db 641 CCGAGGGGTGTTCCGGAGCCCACTGTGGAGTGGGTGCTCTGACAGTCCACAGGGAAGCT 700
Qy 674 GTAAAGAGAAAGTCCAGCTGTTGTTAAAAAGAGGAGAAAGAGTCTTCATGAATTTTG 733
Db 701 GTAAAGAGAGCGCCCTGCTGTGTCAGAAAGAGGAGAAAGGATCTTCATGAGTGTTCG 760
Qy 734 GGACGACATAAAGTGTCTGCTGCCAGAAATGAATGGGCGAGGGAATGCACAGGCTGTTC 793
Db 761 GAACAGACATCAGATCTGCTGTAGAAATGCACCTGGCGCGCGAATCGACCAAGCTGTTC 820
Qy 794 CAATAGATCAATCAAACTCTCAGACACATTCAGACCAATTTTCTTAAAGTAGGGG 853
Db 821 CCATAGATCAAAACAGGCTCTCAGACACACTGCCCCAGGATTTATCTCTGAAAGTGGGG 880
Qy 854 AACCTTATGGATAAGTGTCAAGCTGTTTATGTGAACCATGGATTCGGCTCACCTGGG 913
Db 881 AACCTTGTGGATCAGGTGTAGGCCATCCATGTGAACCATGGATTCGGCTCACCTGGG 940
Qy 914 AATTAGAAACAAAGACCTCGAGGAGGCACTTCTTTGAGTGAAGTACCTATTCAACAA 973
Db 941 AGCTGGAAGACAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACA 1000
Qy 974 ACAGAACTATGATACCGATTCTGTTTCTTTTGTATCATCAGTGGGAGAAACAGCACCG 1033
Db 1001 ACAGGACCATGATTCGGATTCTCTTGGCTTTTGTCTTCCGTTGGGAGGAAACAGCACCG 1060
Qy 1034 GATACTACACTTCTCTCTTCAAGCATCCAGTCAATCAGCTTTTGGTTACCATCGTAG 1093
Db 1061 GATATTACACTCTCTTCTCTCAAGACCCCGAGCAGTACGCGTTGTTGACCATCTAG 1120
Qy 1094 GAAAGGATTTATAAATGCTTACCAATTTCAAGTGAAGATTTGAATTTGACCAATATGAAG 1153

```

Db	1121	AAAAAGGTTTATTAACCGCTACCAAGCTCGCAAGAGAGTATGAATTGACCCGTACGAAA	1180
QY	1154	AGTTTGTGTTTCTGTAGGTTTAAAGCCTACCCACAAATCAGATGTACGTGGACCTTCT	1213
Db	1181	AGTTCTGCTCTAGTCAGGTTTAAAGGTTACCCACGAATCCGATGCACGTGGATCTTCT	1240
QY	1214	CTCGAAAATCATTTCTTGTGTAGCAAAAGGCTCTTGATACGGATACAGCATATCCAAAGT	1273
Db	1241	CTCAAGCCTCATTTCTTGTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAAT	1300
QY	1274	TTTGCATATCAACCAAGCAGGAGAGATATATATTCATGCAAGAAATGATGCGCC	1333
Db	1301	TTTGCATATCAAGCAAGCAGGAGAGATATATATTCATGCAAGAAATGATGCGCC	1360
QY	1334	AAATTACCAAAATCTTCACGCTGAATATAGCAAGAAACCTCAAGTCTCCCAAGAGCAT	1393
Db	1361	AGTTCACCAAAATGTTTCACGCTGAATATAGCAAGAAACCTCAAGTCTCCCAAGAGCAT	1420
QY	1394	CGCAAGCTCAGGCGTCTGTTCTCGGATGGATACCCATTAACCATCTTGGACCTGGGAAG	1453
Db	1421	CAGCCAGCCAGGCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGGAAG	1480
QY	1454	AGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATACAGAAAGGAGTCTGGAAATAGAA	1513
Db	1481	AGTGTTCGGCAAAATCTCCCAATTTGCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAA	1540
QY	1514	AGCTTAACAGAAAGTCTTGGACAGTGGGTGTCGACGAGTACTCTAAACATGAGTGAAG	1573
Db	1541	AGCTTAACAGAAAGTCTTGGCCAGTGGGTGTCGACGAGTACTCTAAATATGAGTGAAG	1600
QY	1574	CCATAAAGGTTCTCTGCTCAAGTGTCTGCATACAAATCCCTTGGCACATCTTTGTGAGA	1633
Db	1601	CCGGAAAGGCTCTGCTCAAAATGCTGTGGTACAAATCTATGCGGACGCTCTTGGCAAA	1660
QY	1634	CGATCTCTTTAACTCTCCAGGCCCCCTCCCTTTTCATCCAGACAACATCTCATTTCTATG	1693
Db	1661	CCATCTTTTAAACTCAACAGGCCCCCTCCCTTTTCATCCAGACAACATCTCATTTCTATG	1720
QY	1694	CAACAATGTGTGTTGCTCTCTTCAATGCTGCTTTTAAACCTGCTAATTTGTGCACAAGT	1753
Db	1721	CGACCAATGGGCTGTGTCTCCCTTCAATGCTGCTTTTAAACCTGCTAATTTGTGCACAAGT	1780
QY	1754	ACAAAAGCAATTTAGGTATGAAGCCAGGTACAGATGGTACAGGTGACCGGCTCCCTCAG	1813
Db	1781	ACAAAAGCAATTTAGGTACGAGTCACTGTCAGATGATCCAGGTGACGCCCCCTGG	1840
QY	1814	ATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTCAAAATGGGAGTTTC	1873
Db	1841	ATAACGAGTACTTCTACGTTGACTTTCAGGAGTATGAATATGACCTTTAAGTGGGAGTTTC	1900
QY	1874	CAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGATCAGGTGCTTTTGGAAAAGTGATGA	1933
Db	1901	CGAGAGAGAAATTTAGAGTTTGGGAAGGTCTCGGCTCGGCTTTTCGGGAGGTGATGA	1960
QY	1934	ACGCAACAGCTTTAGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTTGCGCTCAAAATGC	1993
Db	1961	ACGCCAGGCTATGCGATTTAGTAAACGGAGTCTCAATTCAGGTCGGGCTGAAGATGC	2020
QY	1994	TGAAGAAAAGCAGACAGCTCTGAAGAGAGGACCTCATGTGACAGAACTCAAGATGATGA	2053
Db	2021	TAAAAGAAAAGCTGACAGCTGTGAAAAACAAGCTCTCATGTGCGAGCTCAAAATGATGA	2080
QY	2054	CCAGCTGGGAAGCCAGGAGAAATTTGTGAACCTGCTGGGGGCTGACACATGTCAGGAC	2113
Db	2081	CCACCTGGGACACCATGACAACATCGTGAATCTGTGGGGGATGACACACTGTCAGGCG	2140
QY	2114	CAATTACTTGTATTTTGTGATCTGTTGCTATGGTATCTTCAACTATCTAAGAGTA	2173
Db	2141	CAGTGTACTTGTATTTTGTGAATATTTGTTGCTATGGTGACCTCTCAACTTCAAGAGTA	2200
QY	2174	AAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACACAATTTTCAAGTTT	2233
Db	2201	AAAGAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACACAATTTTCAAGTTT	2260
QY	2234	ACCCACCTTTTCCAAATCACATCCAAATTTCCAGCATCGCTGGTTCAAGAGAGTTTCAGATAC	2293
Db	2261	ACCCCTACTTTTCCAGCAATTTCAAAATTTCCAGCATCGCTGGTTCCAGAGAGTTTCAGTAC	2320
QY	2294	ACCCGAGCTCGGATCAAAATCTCAGGCTTCAATGGGAATTCATTTCACTCTGAAGATGAAA	2353
Db	2321	ACCCGCCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGA	2380
QY	2354	TTGAATATGAAGAACAAAAAGGCTG-----GAAGAGAGAGGAGACTTTGAATGTCTTA	2407
Db	2381	TTGAATATGAAGAACAAAAAGGCTGCGAGAAAGAGAGGAGAGATTTTGAAGCTCTGA	2440
QY	2408	CATTTGAAGATCTCTTTTGTCTTGGATATCAAGATTGCCAAAGAAATGGAATTTCTGGAAT	2467
Db	2441	CGTTTGAAGAGCTCTCTTTTGTCTTGGATATCAAGATTGCCAAAGAAATGGAATTTCTGGAAT	2500
QY	2468	TTAAGTCTGTGTTTCCAGAGACCTTGGCGCCAGGAACGCTGTGTACCCACGGGAAAG	2527
Db	2501	TCAAGTCTGTGTTTCCAGAGACCTTGGCGCCAGGAATGTGTGTTACCCACGGGAAAG	2560
QY	2528	TGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGAATTTCCAACTATGTTG	2587
Db	2561	TGGTGAAGATCTGTGACTTTTGGACTTGGCGCCAGAGATCTCTGAGGACTCCAGCTAGTCG	2620
QY	2588	TCAGGGCAATGCCGCTGTGCTGTGATAAATGGATGGCCCCGAAAGCCGTTTGAAGGCA	2647
Db	2621	TCAGGGCAAGCCAGCGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGCA	2680
QY	2648	TCTACACATTAAGAGTATGCTGCTCATATGGAATATTACTGTGGGAATCTTCTCAC	2707
Db	2681	TCTACACATTAAGAGTATGCTGCTCATATGGAATATTACTGTGGGAATCTTCTCAC	2740
QY	2708	TTGTTGTGAATCTTACCTTGGCATTTCCGTTGTGATGCTAACTTCTACAACTGATTCAAA	2767
Db	2741	TGGTGTGAACCTTTACCTTGGCATTTCCGTTGTGATGCTAACTTCTACAACTGATTCAAA	2800
QY	2768	ATGGATTAAATGGATCAGCAATTTTATGTACAGAGAAATATACATATATGCAAT	2827
Db	2801	GTGGATTAAATGGATCAGCAATTTTATGTACAGAGAAATATACATATATGCAAT	2860
QY	2828	CCTCTGGCTTTTACCTCAAGGAACCGGCTCCTTCCCTAATTTGACTTCGTTTGTAG	2887
Db	2861	CCTCTGGCTTTTACCTCAAGGAACCGGCTCCTTCCCTAATTTGACTTCGTTTGTAG	2920
QY	2888	GATGTCAGTGGCAGATGCAGAAAGCATGTATCAGAAATGTGATGCTGCGCTGTTTCGG	2947
Db	2921	GATGTCAGTGGCAGATGCAGAAAGCATGTATCAGAAATGTGATGCTGCGCTGTTTCGG	2980
QY	2948	AATGCTCTCACCTTACCAAAACAGGCGACCTTTTACGAGAGAGATGGAATTTGGGCTAC	3007
Db	2981	AACATCTCATCTTACCAAAACAGGCGGCTTCCAGAGAGGCGGCTCAGAGCCG	3040
QY	3008	TCCTCTCCGAGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGGACTTCA	3067
Db	3041	CATCGCCACAGGCGGCTTCAAGATTTCAGAGAGAAAGGTTAGCGAGGAGGCTTGA	3099
QY	3068	CCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCACTAA	3127
Db	3100	CCCGCCCA-----CCCTAGCAGGCTGTAGACCGCAGAGCCAAAGATTAGCCTCGCTCT--G	3153
QY	3128	AGAAAATCTATTTATCACTGCTTCCAGGACTTTTCTCTAGAGCGCTCTGGGTTTA	3187
Db	3154	AGGAAGCGCTTACAGGCGGTTGCTTCCGCTGAGCTTTTCTCTAGATGCTGTCTGCAATTA	3213
QY	3188	CTCTTTTCAAGGAGCTTTTGAATCAAAATCAATCTCTGTCTACAGGAGGAGGAGC	3247
Db	3214	CTC-----CAAGAGTACTTCTATAAATCAAACTCTCTCTCGCACAGGCGGAGAGCC	3266
QY	3248	TGATAATGAATTTTGGAGCATTTGATCTCATCCAGGCTTCTCTAGGCGGCTTCTAG	3307
Db	3267	ATAATGAGACTTGTGTTGGTGGAGGCTACCTTGGGGGCTTCTTCCAGGCGGCTTCTAG	3326

QY 3308 TGAATTG-----TGTACTGTAAGTACAGTATATTTCTTTGTAATAATACATAAAACAAA-- 3357
Db 3327 TTGAGGGGAAAGCCATGCTACTGAAATATAGTATATCTTCTGTAATAACGTTGAACAAACC 3386
QY 3358 -----ACATTTTCTTAAGAGAGAGCTAAATATGATTTTTT---AAGTCTATGTTTTTAAAT 3410
Db 3387 AAACCCGCTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAATAATCTATTTTTTAAAT 3446
QY 3411 AATATGTAATTTTTCAGCTATTTTAGTGATATATTTTATGGTGGGAATAAAATTTCTTAC 3470
Db 3447 ACTATGTAACCTTTTTCATCTATTTTAGTGATATATTTTATGGATGGAATAAACTTTCTAC 3506
QY 3471 TACAGAAAAA 3485
Db 3507 TGTAGAAAAA 3521

RESULT 12
AR005211
LOCUS AR005211 3453 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5747651.
ACCESSION AR005211
VERSION AR005211.1 GI:3966090
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka, I.R.
TITLE Antibodies against tyrosine kinase receptor flk-1
JOURNAL Patent: US 5747651-A 1 05-MAY-1998;
FEATURES
Location/Qualifiers
1..3453
/organism="unknown"

BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN
Query Match 63.6%; Score 2227.2; DB 6; Length 3453;
Best Local Similarity 79.6%; Pred No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 35 GGGACCCCGGGCTCCGGAGGCCATGCGCGGTTGGC---GGCGACCGGGCACCGTGC 91
Db 8 GCCTACCGCGCGCTCCGGAGGCCATGCGCGGTTGGCGCAGCGCAGCGCGGGCTGC 67
QY 92 CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG 151
Db 68 TGCTGCTGTTGTTTGTGCAATGATTTCTTGAGACCGTTACAAACCAAGACCTGCCTG 127
QY 152 TGATCAAGTGTGTTTAAATCAATATAGAAACAATGATTCATGCTGCGGAGTCAATCAT 211
Db 128 TGATCAAGTGTGTTTAAATCAATGATGAGAAACAATGCTCATCAGCGGGAAGCCATCAT 187
QY 212 CATATCCCATGGTATCAGAAATCCCGGAGACCTCGGGTGTGCGTTGAGACCCAGAGCT 271
Db 188 CGTACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGGCGCCAGAGTG 247
QY 272 CAGGACAGTGTACGAAGCTCGCGCTGTGGAAGTGATGATCTGCTTCCATCACAATGC 331
Db 248 AAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATCAGCCCTGC 307
QY 332 AAGTCTGCTGATGCCCGGAGCAATTCCTGCTCTGGTCTTTTAAGCACAGCTCC 391
Db 308 AAGTGCAGTCCGCCAGCGGGACCTTTCCTGCTCTGGGTCTTTAAGCACAGCTCC 367
QY 392 TGAATTGCCAGCCACATTTGATTTACAAAACAGAGGAGTCTGTTCCATGTTTGA 451
Db 368 TGGGCTGCCAGCCGACCTTGTATTTACAAACAGAGGAATGTTTCCATGGCCATCTGA 427
QY 452 AAATGACAGAAACCCAGCTGGAGAAATACCTACTTATTTTATTCAGAGTGAAGTACCAAT 511
Db 428 ACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCCATATTCAGAGCGAAGCGCAACT 487

QY 512 ACACAATATTTTACAGTGTAGTATAAGAAATACCCCTGCTTTACACATTAAGAAGACCTT 571
Db 488 ACACAGTACTTTCACAGTGAATGTAGAGATACACAGCTGTATGTGTAAAGAGACCTT 547
QY 572 ACTTTAGAAATAGAAACACGAGGACGCCCTGCTGCTATATCTGAGAGCGTTCCAGAGC 631
Db 548 ACTTTAGAAATAGAAACACGAGGATGCACCTCTGCTATCTCCGAGGTTTCCGAGC 607
QY 632 CGATGCTGGAATGGTGTCTTTCCGATTCACAGGGGAAAGCTGTAAAGAAAGAACTCCAG 691
Db 608 CCACTGTGGAGTGGTGTCTCTGCACCTCCACAGGAAAGCTGTAAAGAAAGAAAGCCCTG 667
QY 692 CTGTTGTTAAAGAGGAGAAAGTCTTTCATGAATTTATTTGGGACGACATTAAGTCT 751
Db 668 CTGTTGTCAGAAAGGAGGAAAGTCTTCTATGAGTTGTTTCGGAACAGACATCAGATGT 727
QY 752 GTGCCAGAAATCACTGGGACGGAATGCACACGAGCTGTTACATAGATCTAAATCAAA 811
Db 728 GTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCAACATAGATCTAAACAGG 787
QY 812 CTCTCTCAGACACATTTGCCACAATTTATTTCTTAAAGTAGGGGAACCTTTATGATAG 871
Db 788 CTCTCTCAGACACACTGCCCCAGTTTATCTCTGAAAGTGGGGAACCTTTGTGGATCAG 847
QY 872 GCAAAGCTGTTTCATGTGAACCATGGATTCGGGCTCACCTGGGAAATTTAGAAACAAAGCAC 931
Db 848 GTAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAGAACAAAGCCC 907
QY 932 TCGAGGAGGGCAACTACTTTGAGATGAGTACCTATTTCACAAACACAGAACTATGATACGGA 991
Db 908 TGGAGGAGGGCAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGACCATGATTCGGA 967
QY 992 TTCTGTTTGTCTTTTCTATCATCAGTGGCAAGAAACACACACCGGATACTACCTTGTCTCT 1051
Db 968 TTCTGTTTGTCTTTTCTCTCCGTTGGGAGGAGACACACCGGATATTACACCTGCTCT 1027
QY 1052 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTAGAAAGGATTTATTAATG 1111
Db 1028 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTAGAAAGGATTTATTAATG 1087
QY 1112 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTAGAAAGGATTTATTAATG 1171
Db 1088 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTAGAAAGGATTTATTAATG 1147
QY 1172 GGTTTAAAGCCTTACCACCAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCCCT 1231
Db 1148 GGTTTAAAGCCTTACCACCAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCCCT 1207
QY 1232 GTGAGCAAAAGGCTTTGATAACGGATACAGCATATCCAAAGTTTTCGAATCATAGCAAC 1291
Db 1208 GTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGCATCATAGCAACA 1267
QY 1292 AGCCAGGAGAAATATATATTTCCATGCAGAAATGATGCCCCAATTTACCAAAATGTTCA 1351
Db 1268 AGCCAGGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTTACCAAAATGTTCA 1327
QY 1352 CGCTGAATATAAGAGAAACCTCAAGTCTCCAGAAAGCTCGCAAGCTCGAGCGCTCC 1411
Db 1328 CGCTGAATATAAGAGAAACCTCAAGTCTCGCAAGCTCGCAAGCTCGAGCGCTCC 1387
QY 1412 GTTCTCTCGGATGATACCCATTTACCATTTGGACCTGGAGAAAGTGTTCAGACAAGTCT 1471
Db 1388 GTTCTCTGATGCTACCCGCTACCCCTTTGGACCTGGAGAAAGTGTTCGGAACAATCTC 1447
QY 1472 CCAACTGCACAGAGATTCACAGAGGAGTCTGGAATAGAAAGGCTAACAGAAAGTGT 1531
Db 1448 CCAATTTGCAGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTTAACAGAAAGTGT 1507
QY 1532 TTGGACAGTGGGTGTCGACGAGTACTTAAACATGAGTGAAGCAATAAAGGTTCTCTCG 1591
Db 1508 TTGGCCAGTGGGTGTCGACGAGTACTTAAATAGTGAAGCGGGAAGGCTTCTCTCG 1567
QY 1592 TCAAGTGTGCTGATACAAATTCCTTTGGCACATCTTTGTGAGACGATCTTTTAACTCTC 1651

```

Db      1568  TCAATGCTGCGTACAAATTTCTATGGGCACGCTCTGGGAAACCATCTTTTAAACTCAC 1627
Qy      1652  CAGGCCCTTCCCTTTTCAAGACACATCTCAATCTATGCAACAATTTGGTTGTC 1711
Db      1628  CAGGCCCTTCCCTTTTCAAGACACATCTCTCTATGCGACCAATTTGGGCTCTGTC 1687
Qy      1712  TCCTCTTCATTTGCTTTTAAACCTTGCTTAATTTGTCACAAGTACAAAAGCAATTTAGGT 1771
Db      1688  TCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCACAATATACAAAAGCAATTTAGGT 1747
Qy      1772  ATGAAAGCCACTACAGATGTACAGGTGACCGGCTCTCTCAAGTAATGAGTACTTTCTACG 1831
Db      1748  ACAGAGTCACTGCAGATGATCCAGGTGACTTGGCCCCCTGGATAACAGTACTTCTAGC 1807
Qy      1832  TTGATTTTCAGAGAATATGAATATGATCTCAATGAGGTGTTCCAAAGAGAAAATTTAGAGT 1891
Db      1808  TTGACTTCAGGAGTATGAATATGACCTTAGTGGGAGTTCCCGAGAGAGAACTTAGAGT 1867
Qy      1892  TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGAATGAACCAACAGCTTATGAA 1951
Db      1868  TTGGGAAGTCTCGGGTCTGGCGTTTCGGGAGGGTGATGAACCCACGGCCTATGGCA 1927
Qy      1952  TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGTCAAAATGCTGAAAGAAAAGCAGACA 2011
Db      1928  TTAGTAAACGGGAGTCTCAATTCAGGTGCGGTGAAGATGCTAAAAGAGAAAGCTGACA 1987
Qy      2012  GCTCTGAAAGAGAGGCACTCATGTCAAGTCAAGATGATGACCCAGCTGGGAAGCCACG 2071
Db      1988  GCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
Qy      2072  AGAATATTGTAACCTGCTGGGGCGTGCACACTGTCAGGACCAATTTACTTTGATTTTG 2131
Db      2048  ACAACATCGTAATCTGCTGGGGCATGCAACACTGTCAGGGCCAGTGTACTTTGATTTTG 2107
Qy      2132  AATACTGTGCTATGGTGATCTCTCAACTATCTAAGAAAGTAAAAGAGAAAATTTTACA 2191
Db      2108  AATATTGTTGCTATGGTGACCTCTCACTCACTTAAGAAAGTAAAAGAGAAAGTTTACA 2167
Qy      2192  GGACTTGGACAGAGATTTTCAAGGAACAAATTTTCAGTTTTTACCCCACTTTTCCAATCAC 2251
Db      2168  GGACATGGACAGAGATTTTAAAGGAACATAATTTTCAGTTCTTACCCCTACTTTCCAGGCAC 2227
Qy      2252  ATCCAAATTCAGCATGCCGTGTTTCAAGAGAGTTTCAGATACACCCGACCTCGATCAAA 2311
Db      2228  ATTCAAAATTCAGCATGCCGTGTTTCAAGAGAGTTTCAGTACACCCGCGCTTGGATCAGC 2287
Qy      2312  TCTCAGGGCTTCATGGGAATTCATTTCACCTGGAAGATGAAATTCATATGAAAACCAAA 2371
Db      2288  TCTCAGGGTTCATGGGAATTCATTCATCTGAAGATGAGATTGAATATGAAAACCAAGA 2347
Qy      2372  AAAGGCTG-----GAAGAGAGAGGAGTGAATGTGCTTACATTTTGAAGATCTTCTTT 2425
Db      2348  AGAGGCTGGCAGAGAAAGAGAGAGAGATTTTGAACGTGCTGACGCTTTGAAGACCTCTTT 2407
Qy      2426  GCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAACTGCTGTGTTTACA 2485
Db      2408  GCTTTGGCTACCAAGTGGCCAAAGGCATGGAATTTCTTGGAGTTTCAAGTCGTGTGTTCCACA 2467
Qy      2486  GAGACCTTGGCCGCGCAGGAACGTGCTTGTCAACCCACGGGAAGTGGTGAAGATATGTGACT 2545
Db      2468  GAGACCTTGGCAGCAGGAATGTTGTTGTCACCCACGGGAGGTGGTGAAGATCTGTGACT 2527
Qy      2546  TTGATTTGGCTCGAGATATCATGAGTGAATTCACCACTATGTTGTCAGGGGCAATGCCGTC 2605
Db      2528  TTGAGCTGGCCGCGAGACATCTGAGCGACTCCAGCTACGTCGTGAGGGCAACGACGGC 2587
Qy      2606  TGCTGTAAATGGATGGCCCCCGAAAGCCTGTTTGAAGGCATCTACACCAATTAAGAGTG 2665
Db      2588  TGCCGGTGAAGTGGATGGCCACCGAGAGCTATTTTGAAGGGATCTACACAATCAAGAGTG 2647
Qy      2666  ATGTCTGTCATATGGAATATTACTGTGGGAATCTTCTCACTTGGTGTGAATCCTTACC 2725

```

```

Db      2648  ACGTCTGGTCTAGGCAATCCCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
Qy      2726  CTGGCATTTCCGGTTGATGCTAACTTCTACAACTGATTCAAAATGATTTAAAATGGATC 2785
Db      2708  CTGGCATTTCTCTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAAATGGAGC 2767
Qy      2786  AGCCATTTTATCTCAGAGAAGAAATACATTAATGAATCCTCTCGGCTTTTGACT 2845
Db      2768  AGCCATTTTATCTCCACAGAAGGATATATTTGTAATGAATCCTCTCGGCTTTTGACT 2827
Qy      2846  CAAGAAAGCGGCATCTTCCCTCAATTTGACTTCGTTTTTATAGGATGTCAGCTGGCAGATG 2905
Db      2828  CAAGGAAGGGCCATCTTCCCCAACCTGACTTCATTTTATAGGATGTCAGCTGGCAGAGG 2887
Qy      2906  CAGAAGAAGCGATGATCAGAATATGGATGGCGGTGTTTCGGAATGTCCTCACACCTTACC 2965
Db      2888  CAGAAGAAG-----CATGTATCAGAACAATCCATCCATCTACTACC 2924
Qy      2966  AAAACAGCGGCCTTTTACAGCAGAGATGGATTTGGGGCTACTCTCTCCGACGGCTCAGG 3025
Db      2925  AAAACAGCGCGCCCTTCAGCAGAGAGCGGGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983
Qy      3026  TCGAAGATTCTGAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
Db      2984  T-GAAGATTACAGAGAAGAAAGTTAGCGAGAGGCTTTGGACCCCGC-----CACCTAG 3038
Qy      3086  CAGGCTGTAGATTACCAAAACAAAGATTAATTTTATCAGTCAATAAGAAAATCTATTATCAAC 3145
Db      3039  CAGGCTGTAGACCGCAGAGCCAAAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACACGG 3096
Qy      3146  TCCTGCTTCAACAGACTTTTCTCTAGACCGCTCTCGGTTTACTCTTTTCAAGGGA 3205
Db      3097  CGTTGCTTTCGTTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAAGTGA 3149
Qy      3206  CTTTGTGTAATCAATCATCTGTGCACAAAGCAGGAGGAGCTGATGAATGAACCTTTATG 3265
Db      3150  CTTCTATAAATCAAACTCTCTCGCACAGCGGGGAGAGCAATAATGAGACTTGTGG 3209
Qy      3266  GAGCATTTGATCTGCATCCAAAGCCTTCTCAGGCGGCTTGTAGTGAATTTGTACCTGAAG 3325
Db      3210  TGAGCCCGCTTACCCTGGGGGCTTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269
Qy      3326  TACAGTATATTTCTTAATAACATAAAACAAA-----AGCATTTTGTGAAGGAGAGC 3378
Db      3270  TATAGTATATTTCTGTAATACTGTAATAAATAATGTAATAATTTTTCAGCTATTAG 3329
Qy      3379  TAATATGATTTTTT--AAGTCTATGTTTAAAATAATATGTAATAATTTTTCAGCTATTAG 3436
Db      3330  TAAATATGATTTTAAAATCTATGTTTAAAATACTATGTAATAATTTTTCATCTATTAG 3389
Qy      3437  TGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAATAAAAAAAAAAAAAA 3496
Db      3390  TGATATATTTTATGATGGAATAAATTTCTACTGTAAAAATAAAAAAAAAAAAAA 3449
Qy      3497  AAAA 3500
Db      3450  AAAA 3453

```

RESULT 13

```

AR071703
LOCUS      AR071703      3453 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5912133.
ACCESSION AR071703
VERSION    AR071703.1 GI:7222591
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 3453)
AUTHORS   Lemischka,I.R.
TITLE      Method for isolating stem cells expressing flk-1 receptors
JOURNAL    Patent: US 5912133-A 1 15-JUN-1999;

```

FEATURES		Location/Qualifiers
source	1..3453	
BASE COUNT	947 a 822 c 875 g 809 t	
ORIGIN	/organism="unknown"	
Query Match	63.6%;	Score 2227.2; DB 6; Length 3453;
Best Local Similarity	79.6%;	Pred. No. 0;
Matches 2775;	Conservative	0; Mismatches 653; Indels 56; Gaps 10;
Qy	35	GGGGACCCGGCTCCGGAGGCCATCGCGGCTGGC---CGCGAGCGGGCACCGCTGC 91
Db	8	GGCTACGCGCGCTCCGGAGGCCATCGCGGCTGGCGCAGCGAGCGCGGGCTGC 67
Qy	92	CGCTGCTGCTGTTTTCGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 151
Db	68	TGCTGCTGTTGTTTGTGCTAGTAATGTTCTGAGACCGTTACAAACCAAGCCTGCTG 127
Qy	152	TGATCAAGTGTGTTTAAATCAATCATAGAACAATGATTCATAGTGGGGAGTCAATCAT 211
Db	128	TGATCAAGTGTGTTTAAATCAATCATAGAACAATGCTCATCAGCGGGAAGCATCAT 187
Qy	212	CATATCCATGGTATCAGAAATCCCGAGACCTCGGCTGCGCTTGAGACCCAGAGCT 271
Db	188	CGTACCGAATGGTCGGAGATCCCGAGAGACCTCCAGTGTACCCGAGGCGCCAGATG 247
Qy	272	CAGGACAGTGTACGAAGCTCGGCTGTGGAAGTGATGTATCTGCTTCCATCACACTGC 331
Db	248	AAGGACCGGTATATGAAGCGCCACCGTGGAGGTGGCGGAGTCTGGTCCATCACCTGC 307
Qy	332	AAGTGTGCTGATGCCCCAGGGAACATTTCTGTCTCTGGGTCTTTAAGCACAGCTCCC 391
Db	308	AAGTGCAGCTCGCCACCCAGGGGACCTTTCTGTGCTCTGGGTCTTTAAGCACAGCTCCC 367
Qy	392	TGAATGCCAGCACATTTTGATTTACAAACAGAGGAGTGTGTTCCATGGTCAATTTGA 451
Db	368	TGGGTGCGCGCGCACTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCTTTGA 427
Qy	452	AAATGACAGAAACCAAGCTGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAATT 511
Db	428	ACGTGACAGAGACCCAGGACAGAGAAATACCTACTCCATATTCAGAGCGAACCGCAACT 487
Qy	512	ACACATAATTTTACAGTGAATTAAGAAATACCTGCTTTTACACATTAAGAAGACCTT 571
Db	488	ACACAGTACTGTTCAGTGAATTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547
Qy	572	ACTTTAGAAAATGGAACACAGGAGCCCTGGTCTGCATATCTGAGACGCTTCCAGAGC 631
Db	548	ACTTTAGGAAGATGGAACACAGGAGTCACTGCTCTGCATCTCCGAGGGTGTTCGGAGC 607
Qy	632	CGATCGTGAATGGGTGCTTTGCGATTACAGGGGAAAGCTGTAAAGAGAAAGTCCAG 691
Db	608	CCACTGTGAGTGGGTGCTCTGCAGCTCCACAGGGAAGCTGTAAAGAGAGGCGCTG 667
Qy	692	CTGTTCTTAAAAGGAGGAAAAGTCTTCATGAATTTATTTGGAGCGGACATTAAGTGTCT 751
Db	668	CTGTTCTGAAAGGAGGAAAAGTACTTCATGAGTTGTCGGAACAGACATCAGATGCT 727
Qy	752	GTGCCAGAAATGAATGCGGACGGGAATGCACAGGCTGTTCAATAGATCTAAATCAA 811
Db	728	GTGCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTCACATAGATCTAAACCCAG 787
Qy	812	CTCCTCAGACCACATTTGCCACAATTTTCTTAAAGTAGGGGAACTCTATGATAGGT 871
Db	788	CTCCTCAGACACACTGCCCGAGTTATTCCTGAAAGTGGGGGAACTCTGTGATCAGGT 847
Qy	872	GCAAGCTGTTATGTGAACCATGGATTCCGGGCTCACCTGGGAAATTAAGAAAACAAAGCAC 931
Db	848	GTAAGGCCATCCATGTGAACCATGGATTCCGGCTCACCTGGGAGTGGAGACAAAGCC 907
Qy	932	TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTTCACAAACAGAACTATGATACGGA 991
Db	908	TGAGGAGGCGACTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967

Qy	992	TTCTGTTTGTCTTTTGTATCATCAGTGGCAAGAAAGACACCGGATACTACACTTGTCTCT 1051
Db	968	TTCTCTTGGCCTTTGTCTTCCGTGGAGAGACGACCGGATATTACACCTGCTCTT 1027
Qy	1052	CTTCAAGACATCCCGAGTCAATCAGCTTTGGTTTACCATCGTAGGAAAGGATTTATAATG 1111
Db	1028	CCTCAAGACACCCCGAGTGGTGGTGGACCATCTAGAAAAAGGTTTATAAAGC 1087
Qy	1112	CTACCAATTCAAGTGAAGATTATGAATTTGACCAATATGAAGAGTTTGTCTTCTGTCA 1171
Db	1088	CTACCAAGTCCGCAAGAAGAGTATGAATTTGACCCGTACGAAAAGTTTGTCTTCTCAGTCA 1147
Qy	1172	GGTTTAAAGCCTACCCCAAAATCAGATGATAGTGGACCTTCTCTCGAAAATCATTTCCCT 1231
Db	1148	GGTTTAAAGCCTACCCCAAGATCCCGATGACAGTGGATCTTCTCTCAAGCCTCATTTCCCT 1207
Qy	1232	GTGAGCAAAAGGCTTGTGATAACGGATACAGCATATCCAAAGTTTGGCAATCATAGCAAC 1291
Db	1208	GTGAACAGAGAGGCTCGGAGGATGGTACAGCATATCTAAATTTTGGCATCATAGAACA 1267
Qy	1292	AGCCAGGAGATATATATTTCCATGAGAAAATGATGCCCAATTTACCAAAATCTTCA 1351
Db	1268	AGCCAGGAGATACATATTTCTATGAGAAAATGATGACCCAGTTTCAACAAATGTTCA 1327
Qy	1352	CGCTGAATATAGAAGGAACCTCAAGTCTCGCAAGACATCGGCAAGTTCAGGCGTCTCT 1411
Db	1328	CGCTGAATATAGAAGGAACCTCAAGTCTAGCAATGCTAGCAATGCCTCAGCCAGCGCTCT 1387
Qy	1412	GTTTCTCGATGATACCATTTACCATTTGGACCTTGGAGAGTGTTCAGACAAGTCTC 1471
Db	1388	GTTTCTCTGATGCTACCGCTTACCTCTTGGACCTTGGAGAGTGTTCGGACAATCTC 1447
Qy	1472	CCAAGTGCACAGAAAGATCAGAGAGTGTGGAATAGAAGGCTTAACAGAAAAGTGT 1531
Db	1448	CCAATTGCACGAGGAAATCCCAAGAGAGTTTGGAAATAAAAGGCTTAACAGAAAAGTGT 1507
Qy	1532	TTGGACAGTGGTGTGCGACAGTACTCTAAACATGATGAGTGAAGCATAAAAGGTTCTCTGG 1591
Db	1508	TTGGCCAGTGGTGTGCGACAGTACTCTAAATATGATGAGTGAAGCCGGGAAAGGCTTCTGG 1567
Qy	1592	TCAAGTGTGTGATACAATTTCCCTTTGGCACATCTTTGTGAGACGATCTTTTAAACTCTC 1651
Db	1568	TCAATGTGTGGTGTGATCTATGGCACGCTTTGGCAACCATCTTTTAAACTCAC 1627
Qy	1652	CAGGCCCCCTTCCCTTTTCATCCAAAGAACATCTCTATTTATGAACAATTTGGTGTGTC 1711
Db	1628	CAGGCCCCCTTCCCTTTTCATCCAAAGAACATCTCTCTTATGCGACCATTTGGGCTCTGTC 1687
Qy	1712	TCTCTTCAATGTGCTTTTAAACCTTCTAATTTGTCACAAGTACAAAAGCAATTTAGGT 1771
Db	1688	TCCCTTCAATGTGCTTTTAAACCTTCTAATTTGTCACAAGTACAAAAGCAATTTAGGT 1747
Qy	1772	ATCAAGACAGCTACAGATGTTACAGTGACCGGCTCTCTCAGATAATGATGACTTCTACG 1831
Db	1748	ACGAGAGTCACTGTCAGATGATCCAGTGACTGCGCCCTGGATAACGAGTACTTCTACG 1807
Qy	1832	TTGATTTTCAAGAAATATGAATATGATCAAAATGGGAGTTTCCAAAGAGAAAATTTAGAGT 1891
Db	1808	TTGATTTTCAAGAAATATGAATATGATCAAAATGGGAGTTTCCAAAGAGAAAATTTAGAGT 1867
Qy	1892	TTGGGAAGGTACTAGGATCAGGTCTTTTGGAAAAGTGAAGCAACAGCTTATGGAA 1951
Db	1868	TTGGGAAGGTCTGGGGCTTGGCGCTTTCGGGAGGTTGATGAAGCCACGCGCTATGGCA 1927
Qy	1952	TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGTCAAAATGCTGAAAAGAAAAGCAGACA 2011
Db	1928	TTAGTAAAACAGGAGTCTCAATTCAGGTGGCGGTGAGAGTCTTAAGAGAGAAAGCTGACA 1987
Qy	2012	GCTCTGAAAGAGAGCCTCATGTGCGGAGCTCAAAATGATGACCCACCTGGGAGACCCACG 2071
Db	1988	GCTCTGAAAGAGAGCCTCATGTGCGGAGCTCAAAATGATGACCCACCTGGGAGACCCATG 2047


```
QY 2072 AGAATATTGTGAACCTGCGGGCGTGCACACTGTCAAGACCAATTTACTTGATTTTG 2131
Db 2048 ACAACATCTGAATCTGCTGGGGCATGCACACTGTCAAGGCCAGTCTACTTCAATTTTG 2107
QY 2132 AATACTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAAATTCACA 2191
Db 2108 AATATGTTGCTATGCTGATGCTCTCACTCACTCACTCACTCACTCACTCACTCACTCA 2167
QY 2192 GGACTTGGACAGAGATTTTCAAGGAACACAATTTCAAGTTTACCCCACTTTCCATCAC 2251
Db 2168 GGACATGGACAGAGATTTTCAAGGAACACAATTTCAAGTTTACCCCACTTTCCAGGCAC 2227
QY 2252 ATCCAAATTCAGCATGCTGCTGCTCAAGAGAAAGTTCAGATACACCCGGACTCGGATCAAA 2311
Db 2228 ATTCAAATTCAGCATGCTGCTGCTCAAGAGAAAGTTCAGATACACCCGGACTCGGATCACG 2287
QY 2312 TCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAATTTGAAGAAACAA 2371
Db 2288 TCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAATTTGAAGAAACAA 2347
QY 2372 AAAGCTG-----GAAGAGAGAGGAGGACTTGAATGCTGCTTACATTTGAAGATCTCTTT 2425
Db 2348 AGAGCTGGCAGAGAGAGGAGGAGATTTGAAGCTGCTGAGTTTGAAGAGCTCTCTTT 2407
QY 2426 GCTTTCATATCAAGTTGCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2485
Db 2408 GCTTTCATATCAAGTTGCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2467
QY 2486 GAGACCTGGCCCGCAGGACGCTGTTGTCACCCAGGGAAGTGGTGAAGATATGATGACT 2545
Db 2468 GAGACCTGGCCCGCAGGACGCTGTTGTCACCCAGGGAAGTGGTGAAGATATGATGACT 2527
QY 2546 TTGGATTTGGCTGAGATATCATGAGTATCTCAACTGTTGTCAGGGGCAATGCCGCTC 2605
Db 2528 TTGGATTTGGCTGAGATATCATGAGTATCTCAACTGTTGTCAGGGGCAATGCCGCTC 2587
QY 2606 TGCTGTAAATGGAATGGCCCGCAGGAGCTGTTTGAAGGCACTACACCATTAAGAGTG 2665
Db 2588 TGCCGCTGAGTGGATGGCCCGCAGGAGCTTATTTGAAGGCACTACACCATTAAGAGTG 2647
QY 2666 ATGCTGTGCTATATGGAATATTTACTGTGGGAATCTTCTCACTTGGTGTGAATCTTACC 2725
Db 2648 ACCTGTGCTTCTACGGCATCTCTCTCTGGAGATATTTTCACTGGTGTGAACCTTACC 2707
QY 2726 CTGGCATTCGGTGTGATCTACTTCTACAACTGATTCAAATGGAATTAATGGATC 2785
Db 2708 CTGGCATTCGGTGTGATCTACTTCTACAACTGATTCAAATGGAATTAATGGATC 2767
QY 2786 AGCCATTTATGCTACAGAGAAATATACATTAATGCAATCTCTGCTGGCTTTTGACT 2845
Db 2768 AGCCATTTATGCTACAGAGAAATATACATTAATGCAATCTCTGCTGGCTTTTGACT 2827
QY 2846 CAAGAAACGGCCATCTTCCCTAATTTGACTTTTGAAGTGTCAAGTGGCAGATG 2905
Db 2828 CAAGAAACGGCCATCTTCCCTAATTTGACTTTTGAAGTGTCAAGTGGCAGATG 2887
QY 2906 CAGAAAGACGATGATCAGATGAGGATGGCGCTGTTTCGGAATGCTCTACACCTACC 2965
Db 2888 CAGAAAG-----CATGTATCAGAAATCCATCCATCTACC 2924
QY 2966 AAAACAGCGGACCTTTTCAGAGAGAGATGGATTTTGGGGCTACTCTCTCCGAGGCTCAGG 3025
Db 2925 AAAACAGCGGCGCTTCAGAGAGAGGCGGCTCAGAG--CAAGTCGCCACAGCGCGCAGG 2983
QY 3026 TCGAAGATTCGTAGAGAACAAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
Db 2984 T-GAAGATTCAGAGAGAAAGATTTAGCGAGGAGCGCTTGGACCCCGC-----CACCTAG 3038
QY 3086 CAGGCTGTAGATTACCAACAAAGATTAATTTATCATCTATAAAGAAAAATCTATTATCAAC 3145
Db 3039 CAGGCTGTAGACCGCAGAGCAAGATTAGCCTTCGCTCT--GAGGAAGCGCCCTACAGCG 3096
QY 3146 TGCTGCTTACCAGACTTTTCTCTAGAGCGCTGCTGCGTTTACTCTTGTTTTCAAGGGA 3205
```

```
Db 3097 CGTTGCTTCCGTGGACTTTTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA 3149
QY 3206 CTTTGTGTAATAAATCAATCATCCTGTGCACAGGAGGAGGAGCTGATATGAACATTTATTG 3265
Db 3150 CTTTGTGTAATAAATCAACCTCTCCTCGCACAGGCGGAGAGCAATAATGAGACTTGTGTGG 3209
QY 3266 GAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTTTCAGGCGGCTTGTGAGTGAATTTGTCTACCTGAG 3325
Db 3210 TGAGCCCGCTACCTCGGGGCTTTCCACGAGCTTGAGGGAAAGCCATGTATCTGAAA 3269
QY 3326 TACAGTATATTTCTTGTAAATACATAAAACAAA-----AGCATTTTGTCTAAGGAGAAGC 3378
Db 3270 TATAGTATATTTCTTGTAAATACGTGAACAAACCAACCCGCTTTTGTCTAAGGAGAAGC 3329
QY 3379 TAATATGATTTT-----AAGTCTATGTTTTAAATAATATGTAATTTTTCAGCTATTATG 3436
Db 3330 TAAATATGATTTTAAATAATCTATGTTTTAAATACTATGTAATTTTTCATCTATTATG 3389
QY 3437 TGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAAAAATAAAAA 3496
Db 3390 TGATATATTTTATGGATGGAAATAAACTTTCTACTGTAAAAAATAAAAAAATAAAAA 3449
QY 3497 AAAA 3500
Db 3450 AAAA 3453

RESULT 14
LOCUS 125169 3453 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 1 from patent US 5548065.
ACCESSION 125169
VERSION 125169.1 GI:1605039
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka, I. R.
TITLE Tyrosine kinase receptor human flk-2-specific antibodies
JOURNAL Patent: US 5548065-A 1 20-AUG-1996;
FEATURES Location/Qualifiers
Source 1. 3453
BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN

Query Match. 63.6%; Score 2227.2; DB 6; Length 3453;
Best Local Similarity 79.8%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 35 GGGGACCGCGGCTCCGAGGCCATCCGCGCTTGGC---GCCGACGCGGCGACCGCTGC 91
Db 8 GGCTACCGCGCTCCGAGGCCATCGGCGCTTGGCGAGCGCAGCGCGCGCTGC 67
QY 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG 151
Db 68 TGTCTGCTTGTGTTTGTGCTAGTAAATGATCTTTCAGACCGTTTACAAACCAAGACCTGC 127
QY 152 TGATCAAGTGTGTTTAAATCAATCAATGAAGACAAATGATTCATCAGTGGGGAAGTCATCAT 211
Db 128 TGATCAAGTGTGTTTAAATCAATGATGAGAAACAAATGGCTCATCAGCGGAAAGCATCAT 187
QY 212 CATATCCCATGTTATCAGAAATCCCGAGACCTCGGCTGCTGCTTTCAGACCCCGCAGAGCT 271
Db 188 CGTACCGAATGTTGCGAGGATCCCGAGAGACCTCCAGATGTAACCGCGCGCCAGAGT 247
QY 272 CAGGAGACGTGTACGAAGCTGCGCTGTGGAAGTGTGATGTTCTGCTTCCATCACACTGC 331
Db 248 AAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCTTCCATCACCTGC 307
QY 332 AAGTGTGCTGATGATGCGCCAGGGAACATTTCTGCTCTGCGGTCTTTAAGACACAGCTCCC 391
```


QY 812 CTCCTCAGACACATTGGCCACAATATTCTTAAAGTAGGGAAACCCCTTATGCAAGGT 871
DB 878 CTCCTCAGACACATTGGCCACAATATTCTTAAAGTAGGGAAACCCCTTATGCAAGGT 847
QY 872 GCAAAGCTGTTTCATGTAACCATGATTCGGGCTCACCTGGGAATAGAAAACAAAGCAC 931
DB 848 GTAAGGCCATCCATGTAACCATGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGGCC 907
QY 932 TCGAGGAGGGCAACTACTTTCAGATGAGTACCTATTCAACAACAGAACTATGATACGGA 991
DB 908 TGGAGGAGGGCACTACTTTCAGATGAGTACCTATTCAACAACAGAACTATGATACGGA 967
QY 992 TTCCTTTTCTTTCATCATCAGTGGCAAGAAACGACACCGGATACACTTGTCTCT 1051
DB 968 TTCCTTTTCTTTCATCATCAGTGGCAAGAAACGACACCGGATATACACTTGTCTCT 1027
QY 1052 CTTCAAAGCATCCCACTGATCAATCAGCTTTGGTTTACCATCGTAGGAAGGGATTTATAATG 1111
DB 1028 CTTCAAAGCATCCCACTGATCAATCAGCTTTGGTTTACCATCGTAGGAAGGGATTTATAATG 1087
QY 1112 CTACCAATTCAGTGAAGATATGAATTTGACCAATATCAACAGTTCCTTTCTGTCA 1171
DB 1088 CTACCAATTCAGTGAAGATATGAATTTGACCAATATCAACAGTTCCTTTCTGTCA 1147
QY 1172 GGTTTAAAGCCCTACCCACAATTCAGATGATGAGTGGACCTTCTCTGAAATCATTTCCCT 1231
DB 1148 GGTTTAAAGCCCTACCCACAATTCAGATGATGAGTGGACCTTCTCTGAAATCATTTCCCT 1207
QY 1232 GTGAGCAAAAGGCTCTGATACGGATACAGCATATCCAAATTTGGCAATATAGACCC 1291
DB 1208 GTGAGCAAAAGGCTCTGATACGGATACAGCATATCCAAATTTGGCAATATAGACCC 1267
QY 1292 AGCCAGGAGATATATATTCATGCGAGAAATGATGATGCCCAATTTACCAAAATGTCA 1351
DB 1268 AGCCAGGAGATATATATTCATGCGAGAAATGATGATGCCCAATTTACCAAAATGTCA 1327
QY 1352 CGCTCAATATAGAAAGAAACCTCAAGTCTCGCAGAAAGCATCGGCAAGTCAAGGCTCT 1411
DB 1328 CGCTCAATATAGAAAGAAACCTCAAGTCTCGCAGAAAGCATCGGCAAGTCAAGGCTCT 1387
QY 1412 GTTCTCGATGATACCAATTTACCATTTGGACCTGGGAAGTGTTCAGACAACTCTC 1471
DB 1388 GTTCTCGATGATACCAATTTACCATTTGGACCTGGGAAGTGTTCAGACAACTCTC 1447
QY 1472 CCAACTGCACAGAGATACAGAAAGGATCTGGAATAGAAAGGCTACAGAAAGTCT 1531
DB 1448 CCAATTTGCAGGAGGAAATCCAGAAAGGATTTGGAATAGAAAGGCTACAGAAAGTCT 1507
QY 1532 TTGGACAGTGGGTGTCGACAGTACTCTAAACATGATGAGTCAAGCATAAAAGGTTCTCTGG 1591
DB 1508 TTGGACAGTGGGTGTCGACAGTACTCTAAACATGATGAGTCAAGCATAAAAGGTTCTCTGG 1567
QY 1592 TCAAGTGTGTGCATACAAATTCCTTGGCAGATCTTTGTGAGAGCATCTCTTTTAAACTCTC 1651
DB 1568 TCAATTTGCTGCGTACAAATTTCTATGGCAGCTCTTTGGCAACCATCTTTTAAACTCAC 1627
QY 1652 CAGGCCCCCTTCCTTTCATCCAGACAACTCTCATTTCTATGCAACAATTTGGTTGTC 1711
DB 1628 CAGGCCCCCTTCCTTTCATCCAGACAACTCTCATTTCTATGCGACCATTTGGGCTCTGTC 1687
QY 1712 TCTCTTCTATGTCGTTTTTAAACCTCTAATTTGTTCACAGTACAAAAGCAATTTAGGT 1771
DB 1688 TCCCTTCTATGTTGTTTCATTTGTTGATCTGCCCAAAATACAAAAGCAATTTAGGT 1747
QY 1772 ATGAAAGCCAGTACAGATGGTACAGGTGACCGGCTCTCAGATAATAGTACTTCTACG 1831
DB 1748 ACAGAGTACAGTACAGATGGTACAGGTGACCGGCTCTCAGATAATAGTACTTCTACG 1807
QY 1832 TTGATTTACAGAAATATGAATATGATCTCAAATGGGAGTTTCCAGAGAAAATTTAGGT 1891
DB 1808 TTGATTTACAGGAGTATGAATATGATCTTAAATGGGAGTTTCCAGAGAAAATTTAGGT 1867
QY 1892 TTGGGAGGTACTAGGTAGTCTTTTGGAAAAGTGTGATGACGCAACAGCTTATGGAA 1951

DB 1868 TTGGGAGGTCTCTGGGCTTTCGGAGGCTGATGAACGCGCACGCGCTATGGCA 1927
QY 1952 TTAGCAAAACAGAGTCTCAATCCAGGTTGCCCTCAAAATGCTGAAGAAAACGACACA 2011
DB 1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAGAAAAGCTGACA 1987
QY 2012 GCTCTGAAAGAGAGGACATCATGTCAAGACTCAAGATGATGACCCAGCTGGGAGCCACG 2071
DB 1988 GCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCAGCTGGACACCATG 2047
QY 2072 AGAATATTGTAACCTGCTGGGGCGGTGCACACTGTCCAGGACCAATTTACTTGTATTTTG 2131
DB 2048 ACAACATCTGTAATCTGCTGGGGCATGCACACTGTCCAGGGCCAGTGTACTTGTATTTTG 2107
QY 2132 AATACTGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGAGAAAATTTTCACA 2191
DB 2108 AATACTGTTGCTATGCTGATGCTTCAACTATCTAAGAAGTAAAGAGAAAATTTTCACA 2167
QY 2192 GGACTTTGGACAGAGATTTTCAAGGAACACAAATTTTACAGTTTTTACCCCACTTTTCAATCAC 2251
DB 2168 GGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTTACCCCTACTTTCCAGGCAC 2227
QY 2252 ATCCAAATTTCCAGCATGCTGTTCAAGAGAAAGTTCAGATACACCCGGACTCGGATCAAA 2311
DB 2228 ATTCAAATTTCCAGCATGCTGTTCAAGAGAAAGTTCAGATACACCCGGACTCGGATCAGC 2287
QY 2312 TCTCAGGGCTTCATGGGAATTCATTTCACTCTCAAGATGAATTTGAATATGAACCAACAA 2371
DB 2288 TCTCAGGGCTTCATGGGAATTCATTTCACTCTCAAGATGAATTTGAATATGAACCAACAA 2347
QY 2372 AAAGGCTG-----GAAGAGAGAGGAGCTTGAATGTCTTACATTTGAAGATCTTCTTT 2425
DB 2348 AGAGCTGGCAGAGAGAGGAGGAGATTTGAACGCTGCTGACGTTTGAAGACCTCTCTTT 2407
QY 2426 GCTTTGCATATCAAGTTGGCAAGAAAGTGAATTTCTGGAATTTTAACTGCTGTGTTTCA 2485
DB 2408 GCTTTGCTGATCAAGTTGGCAAGAAAGTGAATTTCTGGAATTTTAACTGCTGTGTTTCA 2467
QY 2486 GAGACCTGGCCGACAGAAAGCTGTTGTACACCCAGGGAAGTGGTCAAGATATGTGACT 2545
DB 2468 GAGACCTGGCCGACAGAAAGTGTGTTGTACACCCAGGGAAGTGGTCAAGATATGTGACT 2527
QY 2546 TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTGAGGGCAATCCCGCTC 2605
DB 2528 TTGGACTGGCCGACAGATCTGAGGACTCTGAGCTACGTCGTCAGGGGCAACGACGGC 2587
QY 2606 TGCTGTGTAATGATGGCCCGGCAAGAGCTGTTTGAAGGATCTACACCAATTAAGAGTG 2665
DB 2588 TGCCGCTGAAGTGGATGGCAACCCGAGAGCTTATTTGAAGGATCTACACCAATTAAGAGTG 2647
QY 2666 ATGCTGTGCTATGGAATATTACTGTGGGAATCTTCTCACTTGTGTTGTTGTTTACC 2725
DB 2648 ACCTGTGCTTACGCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
QY 2726 CTGGCATTCGGTTGATGCTAACTTCTACAAACTGATTCAAAATGGATTTAAATGGATC 2785
DB 2708 CTGGCATTCCTGTGACGCTTAACTTCTTAACTGATTACAGTGGATTTAAATGGAGC 2767
QY 2786 AGCCATTTTATGCTACAGAAATATACATTTATATGCAATCTGCTGGGCTTTTGTACT 2845
DB 2768 AGCCATTTTATGCTACAGAAATATACATTTATATGCAATCTGCTGGGCTTTTGTACT 2827
QY 2846 CAGAGAAACGGCATCTTCCCTTAATTTGACTTCTGTTTGTAGGATCTCAGCTGGCAGATG 2905
DB 2828 CAGAGAAACGGCATCTTCCCTTAATTTGACTTCTGTTTGTAGGATCTCAGCTGGCAGATG 2887
QY 2906 CAGAGAAACGGCATCTTCCCTTAATTTGACTTCTGTTTGTAGGATCTCAGCTGGCAGATG 2965
DB 2888 CAGAGAAACGGCATCTTCCCTTAATTTGACTTCTGTTTGTAGGATCTCAGCTGGCAGATG 2924
QY 2966 AAAACAGGCGACCTTTTCAGCAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025

```
Db 2925 AAAACAGCGCGCCCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983
Qy 3026 TCGAAGATTCTAGAGGAACAATTAGTTTTAAGGACTTTCCTCCCTCCACCTATCCCTAA 3085
Db 2984 T-GRAGATTACAGAGAAAGAGTTAGCGAGAGGCTTGACCCCGC----CACCCCTAG 3038
Qy 3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTCATCACTAAAGAAAATCTATTATCAAC 3145
Db 3039 CAGGCTGTAGACCGCAGAGCCCAAGATTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Qy 3146 TGCTGTCTCACCAGACTTTTCTCTAGAGCCGCTCGCTTTACTCTGTTTTCAAAGGA 3205
Db 3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149
Qy 3206 CTTTGTAAATCAATCATCTCTGTCAAGGAGGAGGAGCTGATAATGAACCTTTATTG 3265
Db 3150 CTTCTATAAATCAAACTCTCCTCGCACAGCGGGAGAGCCAAATATGAGACTTGTGG 3209
Qy 3266 GAGCATTGATCTGCATCCAAAGCCCTTCTCAGCGCGGCTTGAGTGAATTTGTACTCTGAAG 3325
Db 3210 TGAGCCCGCCTACCTCGGGGCCCTTCCACGAGCTTGAGGGGAAAGCCATGTAICTGAA 3269
Qy 3326 TACAGTATNTCTGTAAATACATAAAACAAA-----AGCATTTGCTAAGGAGAAGC 3378
Db 3270 TATAGTATATCTTGTAAATACGTGAACCAACCAACCCCTTTTTCCTAAGGAAAGC 3329
Qy 3379 TAATATGATTTTTT--AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436
Db 3330 TAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAATTTTTCATCTATTAG 3389
Qy 3437 TGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAAAATAAAAAA 3496
Db 3390 TGATATATTTATGGATGGAAATAAACTTCTACTGTAAAAATAAAAAA 3449
Qy 3497 AAAA 3500
Db 3450 AAAA 3453
```

Search completed: May 25, 2003, 09:19:07
Job time : 8785.17 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 04:06:12 ; Search time 704.832 Seconds
(without alignments)
11185.997 Million cell updates/sec

Title: US-09-919-408-3
Perfect score: 3501
Sequence: 1 CGAGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	16	AAQ81013
2	3501	100.0	3501	16	AAQ79069
3	3501	100.0	3501	17	AAAT38734
4	3501	100.0	3501	18	AAT72117
5	3501	100.0	3501	20	AAAX77515
6	3499.4	100.0	3501	14	AAQ53503
7	3499.4	100.0	3501	14	AAQ40915
8	3497.8	99.9	3501	14	AAQ35250
9	3474.4	99.2	3476	16	AAQ91536
SUMMARIES					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
1	3501	100.0	3501	16	AAQ81013
2	3501	100.0	3501	16	AAQ79069
3	3501	100.0	3501	17	AAAT38734
4	3501	100.0	3501	18	AAT72117
5	3501	100.0	3501	20	AAAX77515
6	3499.4	100.0	3501	14	AAQ53503
7	3499.4	100.0	3501	14	AAQ40915
8	3497.8	99.9	3501	14	AAQ35250
9	3474.4	99.2	3476	16	AAQ91536

10	3470.2	99.1	3475	16	AAQ00802	Human Flk2/Flt3 ty
11	3086.4	88.2	3120	14	AAQ49756	ptk gene LptK25.
12	3083.2	88.1	3120	16	AAQ03096	Protein tyrosine-k
13	2875.8	82.1	2949	19	AAV39041	Human receptor typ
14	2868.4	81.9	2958	19	AAV39040	Human receptor typ
15	2868.4	81.9	2958	19	AAV39042	Human receptor typ
16	2846	81.3	2982	19	AAV39039	Human receptor typ
17	2835.4	81.0	2978	19	AAV39038	Human receptor typ
18	2273	64.9	3521	16	AAQ00801	Flk2/Flt3 tyrosine
19	2272.2	63.6	3453	16	AAQ81012	Flk2 receptor prot
20	2272.2	63.6	3453	17	AAQ38733	Human foetal liver
21	2272.2	63.6	3453	18	AAQ72118	Murine flk-2 recep
22	2272.2	63.6	3453	20	AAQ77514	Murine flk-2 CDNA.
23	2225.6	63.6	3453	14	AAQ53502	Murine flk-2 CDNA.
24	2225.6	63.6	3453	14	AAQ35249	Murine flk-2 codin
25	2225.6	63.6	3453	16	AAQ79068	Mouse flk-2 CDNA.
26	2222.4	63.5	3453	13	AAQ29954	Murine flk-2 CDNA
27	2222.4	63.5	3453	14	AAQ40914	Murine flk-2 CDNA.
28	1545.8	44.2	2247	19	AAV55294	Sequence pMON32390
29	625.4	17.9	1894	15	AAQ54036	Flk-2ws gene. Mus
30	330.4	9.4	332	13	AAQ29955	Extracellular doma
31	324.2	9.3	3992	23	AAQ79666	DNA encoding novel
32	324.2	9.3	3992	24	ABK48105	Human macrophage c
33	322.6	9.2	3069	21	AA244718	Bovine c-Kit bk-1
34	317	9.1	5098	22	AAQ13425	Murine Kit/stem ce
35	313.2	8.9	5084	19	AAV20443	Human c-kit oncoge
36	313.2	8.9	5084	22	AAQ13426	Human Kit/stem cel
37	313.2	8.9	5084	24	ABL64113	Breast cancer rela
38	313.2	8.9	5084	24	ABL68085	Ovary cancer relat
39	308.4	8.8	2919	20	AAV80687	Hampshire porcine
40	308.2	8.8	6390	23	AAQ79665	DNA encoding novel
41	308.2	8.8	6390	23	AAQ84936	DNA encoding novel
42	302.6	8.6	4054	11	AAQ06869	Sequence encodes p
43	302.6	8.6	6378	24	ABL62372	Colon adenocarcino
44	302.6	8.6	6378	24	ABL68519	Kidney cancer rela
45	302.6	8.6	6378	24	ABK35520	Human endometrial

ALIGNMENTS

RESULT 1	
AAQ81013	
ID	AAQ81013 standard; CDNA; 3501 BP.
XX	
AC	AAQ81013;
XX	
DT	18-AUG-1995 (first entry)
XX	
DE	Flk2 receptor protein-tyrosine-kinase CDNA.
XX	
KW	Human Flk2; receptor protein-tyrosine-kinase; primitive.
KW	hematopoietic cell; fetal liver kinase; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	58..3039
FT	/tag= a
FT	/product= Flk2 receptor protein-tyrosine-kinase
FT	58..138
FT	/tag= b
FT	139..3036
FT	/tag= c
XX	
XX	
PN	W09500554-A.
XX	
PD	05-JAN-1995.
XX	
PF	17-JUN-1994; 94WO-US06944.
XX	
PR	18-JUN-1993; 93US-0080244.
PR	21-JUN-1993; 93US-0081508.

1741 ATTTGTCAAGTACAAAAGCAATTTAGTGTATGAAGCCACGTACAGATGTTACAGGTG 1800
1801 ACCGCTCTCAGATAAAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
1801 ACCGCTCTCAGATAAAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
1861 AAATGGGAGTTTCCCAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
1861 AAATGGGAGTTTCCCAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
1921 GGAAAGTGTGAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 GGAAAGTGTGAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 GCCGTCAAAATGCTGAAGAGAAACAGACAGCTCTGAAGAGAGGCACTCATGTGAGAA 2040
1981 GCCGTCAAAATGCTGAAGAGAAACAGACAGCTCTGAAGAGAGGCACTCATGTGAGAA 2040
2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATTGTCAACCTGCTGGGGCGTGC 2100
2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATTGTCAACCTGCTGGGGCGTGC 2100
2101 ACACGTCTCAGGACCAATTTACTTTGATTTTGAATATCTGTTGCTATGGTGTCTTCTCAAC 2160
2101 ACACGTCTCAGGACCAATTTACTTTGATTTTGAATATCTGTTGCTATGGTGTCTTCTCAAC 2160
2161 TATCTAAGAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2221 AATTTTCAGTTTACCCACCTTTCCAAATCACAATCCAAATTCAGAGTGCCTGTTCAAGA 2280
2221 AATTTTCAGTTTACCCACCTTTCCAAATCACAATTCAGAGTGCCTGTTCAAGA 2280
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340
2341 TCTGAAGATGAATTTGAATATGAAGAACCAAAAGCTGGGAAGAGAGAGGAGTGAAT 2400
2341 TCTGAAGATGAATTTGAATATGAAGAACCAAAAGCTGGGAAGAGAGAGGAGTGAAT 2400
2401 GTGCTTACATTTGAAGATCTCTTTCCTTTCATATCAAGTTGCTGCAAGAGGATGAAATTT 2460
2401 GTGCTTACATTTGAAGATCTCTTTCCTTTCATATCAAGTTGCTGCAAGAGGATGAAATTT 2460
2461 CTGGAATTTAAGTCTGTGTTCCACAGAGACCTGGCGCCAGGAAGTGTCTGTCACCCAC 2520
2461 CTGGAATTTAAGTCTGTGTTCCACAGAGACCTGGCGCCAGGAAGTGTCTGTCACCCAC 2520
2521 GGGAAAGTGGGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATCCCAAC 2580
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATCCCAAC 2580
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATG 2640
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATG 2640
2641 GAAAGCATCTACACCATTAAGAGTGTGCTGCTGATATGGAATATTTACTGTGGGAATTC 2700
2641 GAAAGCATCTACACCATTAAGAGTGTGCTGCTGATATGGAATATTTACTGTGGGAATTC 2700
2701 TTCTCATTGTTGTAATCTTACCTGGCATTCGGTTGATGCTAACTTCTACAACATG 2760
2701 TTCTCATTGTTGTAATCTTACCTGGCATTCGGTTGATGCTAACTTCTACAACATG 2760
2761 ATTCAAAATGGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820
2761 ATTCAAAATGGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820
2821 ATGCAATCTCTGGGCTTTTACCTCAAGGAACGCCATCTCTCCCTAATTTGACTTCG 2880

2821 ATGCAATCTCTGGGCTTTTGACTCAAGGAAACGGCCATCTCTCCCTAATTTGACTTCG 2880
2881 TTTTTRAGGATGTCAGCTGGCAGATGTCAGAGAAGCGATGTATCAGATGTGGATGCGCGT 2940
2881 TTTTTRAGGATGTCAGCTGGCAGATGTCAGAGAAGCGATGTATCAGATGTGGATGCGCGT 2940
2941 GTTTCGGAATGTCCTCACACCTACCAACACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
2941 GTTTCGGAATGTCCTCACACCTACCAACACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACAAATTTAGTTTAAAG 3060
3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACAAATTTAGTTTAAAG 3060
3061 ACTTCATCCCTCGACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
3061 ACTTCATCCCTCGACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
3121 CACTAAAGAAATCTATTATCAACTGCTGCTTCCAGAGACTTTTCTCTAGAGCCGCTCT 3180
3121 CACTAAAGAAATCTATTATCAACTGCTGCTTCCAGAGACTTTTCTCTAGAGCCGCTCT 3180
3181 GCGTTTACTCTTGTTCAAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGGCGAG 3240
3181 GCGTTTACTCTTGTTCAAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGGCGAG 3240
3241 GAGGAGCTGATTAATGAATTTTATGGAGCAATGATCTGCATCCAGGCCCTTCTCAGGCCG 3300
3241 GAGGAGCTGATTAATGAATTTTATGGAGCAATGATCTGCATCCAGGCCCTTCTCAGGCCG 3300
3301 GCTTGAGTGAATTTGCTACCTGAAGTACAGTATATTTCTTAAATACATAAAACAAAGC 3360
3301 GCTTGAGTGAATTTGCTACCTGAAGTACAGTATATTTCTTAAATACATAAAACAAAGC 3360
3361 ATTTTCTAAGGAGAAGCTAAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
3361 ATTTTCTAAGGAGAAGCTAAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
3481 AAAAAAAAAAAAAAAAAAAAAA 3501
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 2
AAQ79069
ID AAQ79069 standard; cDNA; 3501 BP.
XX AC AAQ79069;
XX AC
XX XX
DT 04-JUL-1995 (first entry)
XX Human flk-2 cDNA.
DE
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell; ds.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 58..3039
FT /*tag= a
FT sig_peptide 58..138
FT /*tag= b
FT mat_peptide 139..3036
FT /*tag= c
XX
PN US5367057-A.
XX
PD 22-NOV-1994.

XX 02-APR-1991; 91US-0679666.
PF 02-APR-1991; 91US-0679666.
XX 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 19-NOV-1992; 92US-0977451.
PR 30-APR-1993; 93US-0055269.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1995-005894/01.
DR P-PSDB; AAR67536.
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
PT proliferation and/or stimulation of primitive mammalian
PT haematopoietic stem cells in vitro or in vivo.
XX
XX Disclosure; Fig. 2A-1F; 69pp; English.
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAR67536-70,
CC respectively, and the deduced amino acid sequences in AAR67535-37,
CC respectively.
XX
SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 100.0%; Score 3501; DB 16; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCGGCATCCGAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
|||||
DB 1 CGAGCGGCATCCGAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
|||||
QY 61 CGGCGGTTGGCGGCGACCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
|||||
DB 61 CGGCGGTTGGCGGCGACCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
|||||
QY 121 TTGGGACTATTACAAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
|||||
DB 121 TTGGGACTATTACAAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
|||||
QY 181 AACAAATGATTCAGTGGGGAAGTCATCATATCCCATGCTATCAGAAATCCCGGAA 240
|||||
DB 181 AACAAATGATTCAGTGGGGAAGTCATCATATCCCATGCTATCAGAAATCCCGGAA 240
|||||
QY 241 GACCTCGGTGTCGCTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGCTGCGCTGTG 300
|||||
DB 241 GACCTCGGTGTCGCTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGCTGCGCTGTG 300
|||||
QY 301 GAAGTGGATGATTCGCTTCCATACACACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCT 360
|||||
DB 301 GAAGTGGATGATTCGCTTCCATACACACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCT 360
|||||
QY 361 TCCCTGCTCTGGTCTTTAAGCACAGCTCCCTGAATGCGACCATTTTGAATTTACAA 420
|||||
DB 361 TCCCTGCTCTGGTCTTTAAGCACAGCTCCCTGAATGCGACCATTTTGAATTTACAA 420
|||||
QY 421 AACAGAGGAGTGTGTTTCCATGCTGCTTTTGAATAACAGAAACCCCAAGCTGGAGAAATAC 480
|||||
DB 421 AACAGAGGAGTGTGTTTCCATGCTGCTTTTGAATAACAGAAACCCCAAGCTGGAGAAATAC 480
|||||
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATACCAATATGTTTACAGTGAAGTATAAGA 540
|||||
DB 481 CTACTTTTATTCAGAGTGAAGCTACCAATACCAATATGTTTACAGTGAAGTATAAGA 540
|||||
QY 541 AATACCTTCTTACACATTAGAGACCTTACTTTAGAAATAGGAAACACGAGGCGCC 600
|||||

Db 541 AATACCTTCTTACACATTAGAGACCTTACTTTAGAAATAGGAAACACGAGGCGCC 600
|||||
QY 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTTCGATTC 660
|||||
Db 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTTCGATTC 660
|||||
QY 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720
|||||
Db 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720
|||||
QY 721 CATGAATTTATTTGGGACGGACATAGGTGCTGCCAGAAATGAATGCGGAGGAAATGC 780
|||||
Db 721 CATGAATTTATTTGGGACGGACATAGGTGCTGCCAGAAATGAATGCGGAGGAAATGC 780
|||||
QY 781 ACCAGGCTGTTCAATAGATCTAAATCAACCTCTCAGACCACATTTGCCACATTTATTT 840
|||||
Db 781 ACCAGGCTGTTCAATAGATCTAAATCAACCTCTCAGACCACATTTGCCACATTTATTT 840
|||||
QY 841 CTTAAAGTAGGGGAACCCCTTATGATAAGGTGCAAGCTCTTCAATGTGAACCATGGATTC 900
|||||
Db 841 CTTAAAGTAGGGGAACCCCTTATGATAAGGTGCAAGCTCTTCAATGTGAACCATGGATTC 900
|||||
QY 901 GGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
|||||
Db 901 GGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
|||||
QY 961 ACCTATTCAACAAACAGAACTATGATAGGATCTGTTGCTGTTGTTGTTGTTGTTGTTG 1020
|||||
Db 961 ACCTATTCAACAAACAGAACTATGATAGGATCTGTTGCTGTTGTTGTTGTTGTTGTTG 1020
|||||
QY 1021 AGAAGCAGACCGGATACACTGTTCTCTTCAAGCATCCAGCTCAATCAGCTGTTG 1080
|||||
Db 1021 AGAAGCAGACCGGATACACTGTTCTCTTCAAGCATCCAGCTCAATCAGCTGTTG 1080
|||||
QY 1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAAT 1140
|||||
Db 1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAAT 1140
|||||
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCACCAATCAGATGT 1200
|||||
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCACCAATCAGATGT 1200
|||||
QY 1201 ACCTGACCTTCTCTGAAATCATTTCTGAGCAAAAGGCTTGTGATAAGGATAC 1260
|||||
Db 1201 ACCTGACCTTCTCTGAAATCATTTCTGAGCAAAAGGCTTGTGATAAGGATAC 1260
|||||
QY 1261 AGCATATCCAAGTTTTCGAATCATAGCACCCAGGAGGAATATATATTCATGCGAGAA 1320
|||||
Db 1261 AGCATATCCAAGTTTTCGAATCATAGCACCCAGGAGGAATATATATTCATGCGAGAA 1320
|||||
QY 1321 AATGATGATGCCCAATTTTACCAAAATGTTCAAGTGTGATTAAGAGAAACCTCAAGTG 1380
|||||
Db 1321 AATGATGATGCCCAATTTTACCAAAATGTTCAAGTGTGATTAAGAGAAACCTCAAGTG 1380
|||||
QY 1381 CTGCGAGAGCATCGGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
|||||
Db 1381 CTGCGAGAGCATCGGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
|||||
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAAGAGGA 1500
|||||
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAAGAGGA 1500
|||||
QY 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
|||||
Db 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
|||||
QY 1561 AACATGAGTGAAGCCATATAAGGTTTCTGCTCAAGTGTGCTGATACATATTCCTTGGC 1620
|||||
Db 1561 AACATGAGTGAAGCCATATAAGGTTTCTGCTCAAGTGTGCTGATACATATTCCTTGGC 1620
|||||
QY 1621 ACATCTTGTGAGAGCATCTCTTTAAACTCTCCAGGCGCTTCCCTTCAATCACAAGACAAC 1680
|||||

Db	1621	ACATCTTTGTGAGAGATCCTTTTTAACTCTCCAGGCCCTTCCCTTTCAATCCAGACAAC	1680
Qy	1681	ATCTCATCTTATGCAACAATTTGGTGTGTCTCCTCTCTTCAATGTGCTGTTTTAAACCCCTGCTA	1740
Db	1681	ATCTCATCTTATGCAACAATTTGGTGTGTCTCCTCTCTTCAATGTGCTGTTTTAAACCCCTGCTA	1740
Qy	1741	ATTTGTCCAAAGTACAAAAGCAATTTAGTGATGAAGCCAGCTACAGATGGTACAGGTG	1800
Db	1741	ATTTGTCCAAAGTACAAAAGCAATTTAGTGATGAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	1801	ACCGGCTCCTCAGATATGAGTACTTCTAGTGTGATTTCCAGACAATATGAATATGATCTC	1860
Db	1801	ACCGGCTCCTCAGATATGAGTACTTCTAGTGTGATTTCCAGACAATATGAATATGATCTC	1860
Qy	1861	AAATGGGAGTTTCCAAGAGAAAAATTTAGATTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
Db	1861	AAATGGGAGTTTCCAAGAGAAAAATTTAGATTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
Qy	1921	GGAAAAGTGATGAACGCAACAGCTTATGGAAATTTAGCAAAACAGGAGTCTCAATCCAGTTT	1980
Db	1921	GGAAAAGTGATGAACGCAACAGCTTATGGAAATTTAGCAAAACAGGAGTCTCAATCCAGTTT	1980
Qy	1981	GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTGCAGAA	2040
Db	1981	GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTGCAGAA	2040
Qy	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATTTGTGAACCTGCTGGGGGCGTGC	2100
Db	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATTTGTGAACCTGCTGGGGGCGTGC	2100
Qy	2101	ACACTGTCAGGACCAATTTACTTGATTTTTTGAATACGTGTGCTATGGTGATCTTCTCAAC	2160
Db	2101	ACACTGTCAGGACCAATTTACTTGATTTTTTGAATACGTGTGCTATGGTGATCTTCTCAAC	2160
Qy	2161	TATCTTAAGAGTAAAGAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Db	2161	TATCTTAAGAGTAAAGAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Qy	2221	AATTTACAGTTTTTACCOCACCTTTCCAATCACATCCAAATTTCCAGCATGCCGTGGTCAAGA	2280
Db	2221	AATTTACAGTTTTTACCOCACCTTTCCAATCACATCCAAATTTCCAGCATGCCGTGGTCAAGA	2280
Qy	2281	GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Db	2281	GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Qy	2341	TCCTGAAGTCAAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT	2400
Db	2341	TCCTGAAGTCAAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT	2400
Qy	2401	GTGCTTACATTTGAAGATCTTTTGCTTTGTCATATCAAGTTGCCAAGGAATGGAATTT	2460
Db	2401	GTGCTTACATTTGAAGATCTTTTGCTTTGTCATATCAAGTTGCCAAGGAATGGAATTT	2460
Qy	2461	CTGGAATTTAAGTCGTGTTTCACAGAGACCTTGGCCGCCAGGAACGTGCTTGTCAACCCAC	2520
Db	2461	CTGGAATTTAAGTCGTGTTTCACAGAGACCTTGGCCGCCAGGAACGTGCTTGTCAACCCAC	2520
Qy	2521	GGGAAAGTGTGAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Db	2521	GGGAAAGTGTGAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Qy	2581	TATGTTGTCAGGGCAATGCCGTCTGCCCTGTAATAATGGATGGCCCCGAAAGCCCTGTTT	2640
Db	2581	TATGTTGTCAGGGCAATGCCGTCTGCCCTGTAATAATGGATGGCCCCGAAAGCCCTGTTT	2640
Qy	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGCTCATATGGAATTTACTTGTGGGAATC	2700
Db	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGCTCATATGGAATTTACTTGTGGGAATC	2700
Qy	2701	TTCTCACTTGGTGGAATCCTTACCCTGGCAATTCGGGTTGATGCTAACTTCTACAAACTG	2760
Db	2701	TTCTCACTTGGTGGAATCCTTACCCTGGCAATTCGGGTTGATGCTAACTTCTACAAACTG	2760

QY	2761	ATTCAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761		
QY	2821	ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Db	2821		
QY	2821	ATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAAGGGCCATCCTTCCCTAAATTTGACATCG	2880
Db	2821		
QY	2881	ATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAAGGGCCATCCTTCCCTAAATTTGACATCG	2880
Db	2881		
QY	2881	TTTTTACGATGTCAGCTGGGAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
Db	2881		
QY	2941	TTTTTACGATGTCAGCTGGGAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
Db	2941		
QY	2941	GTTTCGGAATGTCCTCACAACCTACCAAAACAGGGGACCTTTTCAGCAGAGAGATGGATTG	3000
Db	2941		
QY	3001	GTTTCGGAATGTCCTCACAACCTACCAAAACAGGGGACCTTTTCAGCAGAGAGATGGATTG	3000
Db	3001		
QY	3001	GGGCTACTCTCTCCGCAAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
Db	3001		
QY	3061	GGGCTACTCTCTCCGCAAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
Db	3061		
QY	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Db	3061		
QY	3121	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Db	3121		
QY	3121	CACATAAAGAAATCTATTATCAACTGCTGCTTCACCAGACTTTTCTCTAGAAGCCGCTCT	3180
Db	3121		
QY	3181	CACATAAAGAAATCTATTATCAACTGCTGCTTCACCAGACTTTTCTCTAGAAGCCGCTCT	3180
Db	3181		
QY	3181	CGGTTTACTCTGTGTTTCAAAGGACATTTTGTGTAATAACAATCATCTGTCACAAGGCAG	3240
Db	3181		
QY	3241	CGGTTTACTCTGTGTTTCAAAGGACATTTTGTGTAATAACAATCATCTGTCACAAGGCAG	3240
Db	3241		
QY	3241	GAGGAGCTGATAATCAACTTTATTGGAGCATTTGATCTGCATCCCAAGGCCTTCTCAGGCCTG	3300
Db	3241		
QY	3301	GAGGAGCTGATAATCAACTTTATTGGAGCATTTGATCTGCATCCCAAGGCCTTCTCAGGCCTG	3300
Db	3301		
QY	3301	GCTTGAGTGAATTCGTACCTGGAAGTACAGTATATCTTGTGTAATAACAATAACAAGC	3360
Db	3301		
QY	3361	GCTTGAGTGAATTCGTACCTGGAAGTACAGTATATCTTGTGTAATAACAATAACAAGC	3360
Db	3361		
QY	3361	ATTGTTGCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTTAAATAAATATGTAAA	3420
Db	3361		
QY	3421	ATTGTTGCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTTAAATAAATATGTAAA	3420
Db	3421		
QY	3421	TTTTTTCAGCTATTGATGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Db	3421		
QY	3481	TTTTTTCAGCTATTGATGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Db	3481		
QY	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481		
QY	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481		
RESULT 3			
AAT38734			
ID	AAT38734 standard; cDNA; 3501 BP.		
XX			
AC	AAT38734;		
XX			
DT	11-DEC-1996 (first entry)		
DE	Murine foetal liver kinase 2 cDNA.		
XX			
KW	Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;		
KW	monoclonal; antibody; extracellular domain; receptor assay;		
KW	haematopoietic stem cell; ligand; stimulation; proliferation;		
KW	differentiation; treatment; anaemia; bone marrow damage;		
KW	cancer chemotherapy; radiation; ds.		
OS	Mus musculus.		
XX			
FF	Key	Location/Qualifiers	

Db 1381 CTCGAGAGCATCGGCAAGTCAGCGCTCCTGTTTCTCGGTGATACCCATTACCATCT 1440
Qy 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACACAAGA 1500
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACACAAGA 1500
Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGTGTGCGAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGTGTGCGAGTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATAAAGGTTCTGTCTCAAGTCTGTGCATCAATTCCTTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGTTCTGTCTCAAGTCTGTGCATCAATTCCTTGGC 1620
Qy 1621 ACATCTGTGAGAGCATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCACAGAAC 1680
Db 1621 ACATCTGTGAGAGCATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCACAGAAC 1680
Qy 1681 ATCTCATTTCTATGAACAATTTGGTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 ATCTCATTTCTATGAACAATTTGGTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 ATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800
Db 1741 ATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800
Qy 1801 ACCGCTCTCAGATATGAGTACTCTACCTGATTTTCAGAGATATGAATATCATCTC 1860
Db 1801 ACCGCTCTCAGATATGAGTACTCTACCTGATTTTCAGAGATATGAATATCATCTC 1860
Qy 1861 AAATGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920
Db 1861 AAATGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920
Qy 1921 GGAAGAGTATGAACCAACAGCTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAGAGTATGAACCAACAGCTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Qy 1981 GCCGTCAAAATGCTGAAGAAAGAACAGACAGCTCTGAAAGAGGAGGCTATGTCAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAGAAAGAACAGACAGCTCTGAAAGAGGAGGCTATGTCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGAACTCTGGGGCGTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGAACTCTGGGGCGTGC 2100
Qy 2101 ACATGTACAGACCAATTTACTTGAATTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160
Db 2101 ACATGTACAGACCAATTTACTTGAATTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAGAGAAATTTTCAGAGGACTTGGACAGAGATTTTCAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAATTTTCAGAGGACTTGGACAGAGATTTTCAGGAACAC 2220
Qy 2221 AATTTCAAGTTTACCCTTCCCAATTCACATCCAAATTCAGATCCCTGGTTCAAGA 2280
Db 2221 AATTTCAAGTTTACCCTTCCCAATTCACATCCAAATTCAGATCCCTGGTTCAAGA 2280
Qy 2281 GAAGTTTCAGATACACCCGAGCTCGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
Db 2281 GAAGTTTCAGATACACCCGAGCTCGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
Qy 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGGAAGAGAGAGGACTTTGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGGAAGAGAGAGGACTTTGAAT 2400
Qy 2401 GTGCTTACATTTGAGATCTTCTTGTCTTGCATATCAAGTTCGCAAGGATGCAATTT 2460
Db 2401 GTGCTTACATTTGAGATCTTCTTGTCTTGCATATCAAGTTCGCAAGGATGCAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTGTTTCAGAGAGACCTGGCCGCCAGGAGCTGCTTGTCAACCAC 2520
Db 2461 CTGGAATTTAAGTCTGTGTTTCAGAGAGACCTGGCCGCCAGGAGCTGCTTGTCAACCAC 2520

RESULT 4
AAT72117
ID AAT72117 standard; cDNA; 3501 BP.

Qy 2521 GGGAAAGTGGTCAAGATATGTGACATTTGGATTGGCTCGAGATATCATGAGTATTCCAAC 2580
Db 2521 GGGAAAGTGGTCAAGATATGTGACATTTGGATTGGCTCGAGATATCATGAGTATTCCAAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640
Qy 2641 GAAGCATCTACACCATTAAGAGTATGATGCTGGTCAATGGAATATTTACTGTGGGAATC 2700
Db 2641 GAAGCATCTACACCATTAAGAGTATGATGCTGGTCAATGGAATATTTACTGTGGGAATC 2700
Qy 2701 TTCTCAGTGGTGTGAATCCTTTACCTGGCATTCGGTTGATGCTAACTTCTACAACCTG 2760
Db 2701 TTCTCAGTGGTGTGAATCCTTTACCTGGCATTCGGTTGATGCTAACTTCTACAACCTG 2760
Qy 2761 ATTTCAAAATGGAATTTAAAATGATGATGCTAACTTCTACAACCTGATGCTAACTTCTACAACCTG 2820
Db 2761 ATTTCAAAATGGAATTTAAAATGATGATGCTAACTTCTACAACCTGATGCTAACTTCTACAACCTG 2820
Qy 2821 ATGCAATCCTGCTGGCTTTTACCTCAAGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880
Db 2821 ATGCAATCCTGCTGGCTTTTACCTCAAGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880
Qy 2881 TTTTGTAGGATGTCAGTGGCAGATGCAAGAAAGCGATGATCAGAAATGTTGGATGCGCGT 2940
Db 2881 TTTTGTAGGATGTCAGTGGCAGATGCAAGAAAGCGATGATCAGAAATGTTGGATGCGCGT 2940
Qy 2941 GTTTCGGAAATGTCCTCACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGAGTTT 3000
Db 2941 GTTTCGGAAATGTCCTCACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGAGTTT 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAACAAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAACAAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Db 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Qy 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTACAGACCTTTTCTAGAGCCGCTCT 3180
Db 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTACAGACCTTTTCTAGAGCCGCTCT 3180
Qy 3181 GGGTTTACTCTGTTTCAAGGGACTTTTGTAAAATCAAAATCATCTGTCAAGAGCAG 3240
Db 3181 GGGTTTACTCTGTTTCAAGGGACTTTTGTAAAATCAAAATCATCTGTCAAGAGCAG 3240
Qy 3241 GAGGAGCTCAATCAACTTTATGGAGCATTCGATCGATCCAGGCTTCTCAGGCCG 3300
Db 3241 GAGGAGCTCAATCAACTTTATGGAGCATTCGATCGATCCAGGCTTCTCAGGCCG 3300
Qy 3301 GCTTCAGTGAATTTGTACCTGAAAGTACAGTATTTCTTGTAAATACATAAAACAAAGC 3360
Db 3301 GCTTCAGTGAATTTGTACCTGAAAGTACAGTATTTCTTGTAAATACATAAAACAAAGC 3360
Qy 3361 ATTTTGTGAAGAGAGCTAAATGATTTTTTAAAGTCTATGTTTTTAAATATATGTAAA 3420
Db 3361 ATTTTGTGAAGAGAGCTAAATGATTTTTTAAAGTCTATGTTTTTAAATATATGTAAA 3420
Qy 3421 TTTTTCAGCTATTAGTGATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTAGTGATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

```

xx AAT72117;
..X 19-AUG-1997 (first entry)
XX Human flk-2 receptor coding sequence.
DE Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
KW pTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;
KW proliferation; differentiation; mammalian; haematopoietic stem cell;
KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 58..3039
FT /*tag= a
FT /product= Human_flk-2_receptor
FT sig_peptide 58..138
FT /*tag= b
FT mat_peptide 139..3036
FT /*tag= c
XX
XX US5621090-A.
XX
XX 15-APR-1997.
XX
XX 02-APR-1991; 91US-0679666.
XX
XX 26-JUN-1992; 92US-0906397.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1997-235228/21.
XX P-PSDB; AAW19873.
XX
XX protein containing the extracellular domain of human flk-2 - used
XX for identification of primitive haematopoietic cell proliferation
XX and differentiation stimulatory ligands, e.g. for treating anaemia
XX
XX Claim 1; Fig 1B; 55pp; English.
XX
XX This sequence encodes the human fetal liver kinase 2 (flk2). flk-2 is
XX a receptor protein tyrosine kinase (pTK) and is important in transducing
XX putative self-renewal signals from the environment. flk-2 is expressed
XX in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
XX and it is thought that flk-2 is expressed in the entire primitive portion
XX of the haematopoietic hierarchy. The invention concerns a recombinant
XX nucleic acid, preferably mRNA, which encodes a protein containing only
XX the extracellular domain of human flk-2 and lacking the flk-2 intra-
XX cellular catalytic domain. The resultant protein represents a soluble
XX form of flk-2 which is used to isolate specific ligands for flk-2. These
XX ligands can be used to stimulate proliferation and/or differentiation of
XX mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
XX treatment of macrocytic or aplastic anaemia or bone marrow damage caused
XX by cancer treatment or radiation.
XX
XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
XX
XX Query Match 100.0%; Score 3501; DB 18; Length 3501;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CGAGGGCGGATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 1 CGAGGGCGGATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60

```

```

QY 61 CCGCGTTGGCGCGGACGCGGCGGACCGCTGCGCTGCTGTTTCTTCTGCAATGATA 120
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 61 CCGCGTTGGCGCGGACGCGGCGGACCGCTGCGCTGCTGTTTCTTCTGCAATGATA 120
XX
QY 121 TTTGGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTGTTTTATCAATCAATAG 180
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 121 TTTGGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTGTTTTATCAATCAATAG 180
XX
QY 181 AACAAATGATTCATCAGTGGGAAGTCATCATATCCCATGATATCAGAAATCCCGGAA 240
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 181 AACAAATGATTCATCAGTGGGAAGTCATCATATCCCATGATATCAGAAATCCCGGAA 240
XX
QY 241 GACCTCGGGTGTGCTTGGAGACCCAGAGCTCAGGACAGTGTACAGAGCTCCCGTGTG 300
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 241 GACCTCGGGTGTGCTTGGAGACCCAGAGCTCAGGACAGTGTACAGAGCTCCCGTGTG 300
XX
QY 301 GAAGTGGATATCTGCTTCCATCACACTGCAAGTGTGCTGCTGATGCCCCAGGACATT 360
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 301 GAAGTGGATATCTGCTTCCATCACACTGCAAGTGTGCTGCTGATGCCCCAGGACATT 360
XX
QY 361 TCCTGCTCTGGGCTTTTAAAGCACAGCTCCCTGAATTCGACGCCACATTTTGAATTACAA 420
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 361 TCCTGCTCTGGGCTTTTAAAGCACAGCTCCCTGAATTCGACGCCACATTTTGAATTACAA 420
XX
QY 421 AACAGAGGAGTGTGTTTCCATGCTATTTTGAATTCACAGAAACCCAGCTGGAGAAATAC 480
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 421 AACAGAGGAGTGTGTTTCCATGCTATTTTGAATTCACAGAAACCCAGCTGGAGAAATAC 480
XX
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTCACAAATTTGTTTACAGTGTAGTAAAGA 540
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTCACAAATTTGTTTACAGTGTAGTAAAGA 540
XX
QY 541 ATATCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGACGCC 600
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 541 ATATCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGACGCC 600
XX
QY 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATGCTGGAATGGTGGTTCGCAATCA 660
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATGCTGGAATGGTGGTTCGCAATCA 660
XX
QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTGCCT 720
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTGCCT 720
XX
QY 721 CATGAATTTTGGGACGGACATTAAGTCTGTGCCAGAAATGAAGTGGGACGGAATGC 780
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 721 CATGAATTTTGGGACGGACATTAAGTCTGTGCCAGAAATGAAGTGGGACGGAATGC 780
XX
QY 781 ACCAGGCTCTTCACAAATAGATCTAAATCAAACTCCCTCAGACCCACATTTGCCACAATTTT 840
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 781 ACCAGGCTCTTCACAAATAGATCTAAATCAAACTCCCTCAGACCCACATTTGCCACAATTTT 840
XX
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAGAAAGCTGTTTCATGTGAACCATGATTC 900
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 841 CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAGAAAGCTGTTTCATGTGAACCATGATTC 900
XX
QY 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960
XX
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTCTGTTTATCATCAGTGGCA 1020
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTCTGTTTATCATCAGTGGCA 1020
XX
QY 1021 AGAAGCAGACCCGGATACACTTGTTCCTTCAAGCATCCAGTCAATCAGCTTTG 1080
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1021 AGAAGCAGACCCGGATACACTTGTTCCTTCAAGCATCCAGTCAATCAGCTTTG 1080
XX
QY 1081 GTTACCATCGTAGGAAGGATTTTAAATGCTACCAATTCAGTGAAGATTTGAAATT 1140
..X |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1081 GTTACCATCGTAGGAAGGATTTTAAATGCTACCAATTCAGTGAAGATTTGAAATT 1140

```

1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
1201 AGGTGACCTTCTCGAAAAATCATTTCCCTGTGAGCAAAAAGGCTTGTGATAACGGATAC 1260
1201 AGGTGACCTTCTCGAAAAATCATTTCCCTGTGAGCAAAAAGGCTTGTGATAACGGATAC 1260
1261 AGCATATCCAAGTTTGTGCAATCATAGACCCAGCCAGGAGATATATATTCATGCGAGAA 1320
1261 AGCATATCCAAGTTTGTGCAATCATAGACCCAGCCAGGAGATATATATTCATGCGAGAA 1320
1321 AATGATGATGCCAAATTTACCAAAATGTTACAGCTGAATATAGAAGAGAAACCTCAAGTG 1380
1321 AATGATGATGCCAAATTTACCAAAATGTTACAGCTGAATATAGAAGAGAAACCTCAAGTG 1380
1381 CTCGAGAAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGATGGATACCATCT 1440
1381 CTCGAGAAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGATGGATACCATCT 1440
1441 TGGACCTGGAAGAGTTTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAAG 1500
1441 TGGACCTGGAAGAGTTTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAAG 1500
1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGACGTGGGTGTCGAGCAGTACTCTA 1560
1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGACGTGGGTGTCGAGCAGTACTCTA 1560
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAAATTCCTTGGC 1620
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAAATTCCTTGGC 1620
1621 ACATCTGTGACAGCATCTTTAACTCTCAGGCCCCCTTCCCTTTTCATCCAAAGACAAC 1680
1621 ACATCTGTGACAGCATCTTTAACTCTCAGGCCCCCTTCCCTTTTCATCCAAAGACAAC 1680
1681 ATCTCATTTATGCAACAAATTTGGTGTGTCCTCTTCATTTGCTGTTTAAACCTGCTA 1740
1681 ATCTCATTTATGCAACAAATTTGGTGTGTCCTCTTCATTTGCTGTTTAAACCTGCTA 1740
1741 ATTTGTCACAAATGACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGTTTACAGGTG 1800
1741 ATTTGTCACAAATGACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGTTTACAGGTG 1800
1801 ACCGCTCTCAGATAATGAGTACTTCTAGTGTGATTCAGAGATATGAAATGATGATC 1860
1801 ACCGCTCTCAGATAATGAGTACTTCTAGTGTGATTCAGAGATATGAAATGATGATC 1860
1861 AAATGGGAGTTTCCAAAGAAAAATTTAGAGTTTGGAGGTTACTAGGATCAGGTGCTTTT 1920
1861 AAATGGGAGTTTCCAAAGAAAAATTTAGAGTTTGGAGGTTACTAGGATCAGGTGCTTTT 1920
1921 GGAAAGTGATGAAGCCAAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 GGAAAGTGATGAAGCCAAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 GCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
1981 GCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
2041 CTCAGATGATCACCCAGCTGGGAAGCCACAGAGATATGTCGACCTGCTGGGGGCTGC 2100
2041 CTCAGATGATCACCCAGCTGGGAAGCCACAGAGATATGTCGACCTGCTGGGGGCTGC 2100
2101 ACATGCTCAGGACCAATTTTACTTGTATTTTGAATACTGTTGATGCTGATCTCTCAAC 2160
2101 ACATGCTCAGGACCAATTTTACTTGTATTTTGAATACTGTTGATGCTGATCTCTCAAC 2160
2161 TATCTAAGAAAGTAAAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAAAGTAAAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2221 AATTTACGTTTTTACCCCACTTTCCAATCACAATCCAAATTCAGCATGCTGTTCAAGA 2280

2221 AATTTACGTTTTTACCCCACTTTCCAATCACAATCCAAATTCAGCATGCTGTTTCAGA 2280
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTCC 2340
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTCC 2340
2341 TCTGAAGATGAATTTGAATATGAACCAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400
2341 TCTGAAGATGAATTTGAATATGAACCAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400
2401 GTGCTTTACATTTCAAGATCTTTGCTTTGTCATATCAAGTTGTCACCAAGGAATGAATTT 2460
2401 GTGCTTTACATTTGCAAGATCTTTGCTTTGTCATATCAAGTTGTCACCAAGGAATGAATTT 2460
2461 CTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCCGCCAGGACGTCGTTGTCAACCCAC 2520
2461 CTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCCGCCAGGACGTCGTTGTCAACCCAC 2520
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGTATCCAAC 2580
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGTATCCAAC 2580
2581 TATGTTGTGAGGGCAATGCCGCTCTGCCCTGTAAAATGAGTGGCCCCGGAAGCCCTGTTT 2640
2581 TATGTTGTGAGGGCAATGCCGCTCTGCCCTGTAAAATGAGTGGCCCCGGAAGCCCTGTTT 2640
2641 GAAGGCATCTACACCAATTAAGAGTGTGCTGGTGCATATGGAATATTAATCTGTGGAAATC 2700
2641 GAAGGCATCTACACCAATTAAGAGTGTGCTGGTGCATATGGAATATTAATCTGTGGAAATC 2700
2701 TTCTCACTGGTGTGAATCCCTTACCCTGGCATTTCCGGTTGATGCTACTTCTTACAACATG 2760
2701 TTCTCACTGGTGTGAATCCCTTACCCTGGCATTTCCGGTTGATGCTACTTCTTACAACATG 2760
2761 ATTCAAATGGATTTAAAATGAGTACGCCATTTATGCTACAGAGAAATATACATTTATA 2820
2761 ATTCAAATGGATTTAAAATGAGTACGCCATTTATGCTACAGAGAAATATACATTTATA 2820
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAAAAGGCGCATCTTCCCTAAATTTGACATTCG 2880
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAAAAGGCGCATCTTCCCTAAATTTGACATTCG 2880
2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAAAGAGCGATGTATCAGAAATGTGGATGGCGGT 2940
2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAAAGAGCGATGTATCAGAAATGTGGATGGCGGT 2940
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTTAAGG 3060
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTTAAGG 3060
3061 ACTTCACTCCCTCAACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
3061 ACTTCACTCCCTCAACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
3121 CACTAAAGAAATCTATTATCAACTGCTGCTTTCACAGACCTTTTCTCTAGAAGCCGCTCT 3180
3121 CACTAAAGAAATCTATTATCAACTGCTGCTTTCACAGACCTTTTCTCTAGAAGCCGCTCT 3180
3181 GGGTTTTACTTCTGTTTTTCAAGGGACTTTTGTAAAATCAATCATCTGTCACAAAGGCAG 3240
3181 GGGTTTTACTTCTGTTTTTCAAGGGACTTTTGTAAAATCAATCATCTGTCACAAAGGCAG 3240
3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCAATCCAAAGCCCTTCTCAGGCCG 3300
3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCAATCCAAAGCCCTTCTCAGGCCG 3300
3301 GCTTGAGTGAATGTGTACTGAAGTACAGTATATCTTGTGTAATACATAAAACAAGC 3360
3301 GCTTGAGTGAATGTGTACTGAAGTACAGTATATCTTGTGTAATACATAAAACAAGC 3360

3301 GCTTCAGTGAATGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
3361 ATTTTGTCTAAGCAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATAATATGATAA 3420
3361 ATTTTGTCTAAGCAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATAATATGATAA 3420
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480
3481 AAAAAAAAAAAAAAAAAAAAAA 3501
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 5
AA77515
ID AAX77515 standard; cDNA; 3501 BP.
XX AAX77515;
XX 05-AUG-1999 (first entry)
XX Human flk-2 cDNA.
XX Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
XX monoclonal; polyclonal; antibody; tyrosine kinase; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 58..3039
XX FT /*tag= a
XX FT /product= "flk-2"

US912133-A.
15-JUN-1999.
10-FEB-1998; 98US-0021324.
19-NOV-1992; 92US-0977451.
02-APR-1991; 91US-0679666.
28-JUN-1991; 91US-0728913.
15-NOV-1991; 91US-0793065.
24-DEC-1991; 91US-0813593.
26-JUN-1992; 92US-0906397.
12-NOV-1992; 92US-0975049.
30-APR-1993; 93US-0055269.
31-OCT-1994; 94US-0252498.
15-FEB-1996; 96US-0601891.
(UYPR-) UNIV PRINCETON.
Lemischka IR;
WPI; 1999-357194/30.
P-PSDB; AAY08617.

Isolating hematopoietic cells expressing fetal liver kinase 1
receptors
Disclosure; Fig 1b; 59pp; English.
This invention describes a novel method of isolating cells expressing fetal liver kinase 1 (flk-1) receptors on their surface and comprises binding the cells to a polyclonal or monoclonal antibody specific to the Flk-1 receptor and isolating the cells that have bound to the antibody. The method can be used to isolate hematopoietic stem cells in any mammal but preferably a rat, mouse, rabbit or human. The proteins of the invention belong to the receptor protein family. This sequence encodes the human flk-2 protein which is used in the method of the invention.

XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
SQ
Query Match 100.0%; Score 3501; DB 20; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGGCGGATCGGAGGGCTGGCGGGCCCTGGGGACCCCGGGCTCCGAGGGCCATG 60
DB 1 CGAGGGCGGATCGGAGGGCTGGCGGGCCCTGGGGACCCCGGGCTCCGAGGGCCATG 60
QY 61 CCGGCGTTGCGCGCGACGCGGCGACCGTGCCTGCTGCTGTTTCTGCAATGATA 120
DB 61 CCGGCGTTGCGCGCGACGCGGCGACCGTGCCTGCTGCTGTTTCTGCAATGATA 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATATAAG 180
DB 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATATAAG 180
QY 181 AACATGATTTCATCAGTGGGAAAGTCAATCATATCCCATGTTATCAGAAATCCCGGAA 240
DB 181 AACATGATTTCATCAGTGGGAAAGTCAATCATATCCCATGTTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGTGTGCGTTGAGACCCAGAGCTCAGGGACGTGTACGAAGCTGCCGCTGTG 300
DB 241 GACCTCGGTGTGCGTTGAGACCCAGAGCTCAGGGACGTGTACGAAGCTGCCGCTGTG 300
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTCAAGTGTGTTGCGATGCCCGAGGAACATT 360
DB 301 GAAGTGGATGTATCTGCTTCCATCACACTCAAGTGTGTTGCGATGCCCGAGGAACATT 360
QY 361 TCCTGTCTCGGTGTTTAAAGCACAGCTCCCTGAATTCGACGACCATTTTGAATTTACAA 420
DB 361 TCCTGTCTCGGTGTTTAAAGCACAGCTCCCTGAATTCGACGACCATTTTGAATTTACAA 420
QY 421 AACAGAGGAGTTGTTCCATGGTTCATTTGAARAATGACAGAAACCCAGCTGGAGAAATAC 480
DB 421 AACAGAGGAGTTGTTCCATGGTTCATTTGAARAATGACAGAAACCCAGCTGGAGAAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGCTTACCAATTTACAAATATTTTACAGTGAGTATAAGA 540
DB 481 CTACTTTTATTCAGAGTGAAGCTTACCAATTTACAAATATTTTACAGTGAGTATAAGA 540
QY 541 AATACCTGCTTTACACATTAAAGAACCTTTACTTTTGAANAATGGAACACAGACGCC 600
DB 541 AATACCTGCTTTACACATTAAAGAACCTTTACTTTTGAANAATGGAACACAGACGCC 600
QY 601 CTGCTCTGCATATCTGAGAGCGGTTCCAGAGCCGATCGTGGAAATGGTCTTTGGGATTC 660
DB 601 CTGCTCTGCATATCTGAGAGCGGTTCCAGAGCCGATCGTGGAAATGGTCTTTGGGATTC 660
QY 661 CAGGGGAAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTGCTT 720
DB 661 CAGGGGAAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTGCTT 720
QY 721 CATGAATTTATTTGGGACGGACATTAAGTGTGTTGCCAGAAATGAACCTGGGAGGAATGC 780
DB 721 CATGAATTTATTTGGGACGGACATTAAGTGTGTTGCCAGAAATGAACCTGGGAGGAATGC 780
QY 781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACACATTCGCCCAATATT 840
DB 781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACACATTCGCCCAATATT 840
QY 841 CTTAAAGTAGGGGAACCTTTATGGATGAAGTGAAGAGCTGTTTATGTGAACCATCGAATTC 900
DB 841 CTTAAAGTAGGGGAACCTTTATGGATGAAGTGAAGAGCTGTTTATGTGAACCATCGAATTC 900
QY 901 GGGCTCACCTGGGAATTAAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
DB 901 GGGCTCACCTGGGAATTAAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
QY 961 ACCTATTTCACAAAACAGAACTATGATACGGATTCGTTTGTGTTTATCATCATCAGTGGA 1020
DB 961 ACCTATTTCACAAAACAGAACTATGATACGGATTCGTTTGTGTTTATCATCATCAGTGGA 1020

Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGGCTTTGTATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATCTACACATGTTCCTCTCAAGACATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATCTACACATGTTCCTCTCAAGACATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATTCTGAGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
Db 1081 GTTACCATTCTGAGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
Qy 1141 GACCAATATGAAGAGTTTGTTCCTGTCAGGTTTAAAGCTCCCAACAATCAGATGT 1200
Db 1141 GACCAATATGAAGAGTTTGTTCCTGTCAGGTTTAAAGCTCCCAACAATCAGATGT 1200
Qy 1201 ACCTGGACCTCTCTCGAAATATCTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260
Db 1201 ACCTGGACCTCTCTCGAAATATCTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260
Qy 1261 AGCATATCCAAAGTTTGCATATGAAGCACCAGCAGGAGAAATATATATTCATTCAGAA 1320
Db 1261 AGCATATCCAAAGTTTGCATATGAAGCACCAGCAGGAGAAATATATATTCATTCAGAA 1320
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGGTGAATATAGAAGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGGTGAATATAGAAGAAACCTCAAGTG 1380
Qy 1381 CTCGCAAGCATCGGCAAGTCAAGGCTCTCTGTTCTCGGATGATACCCATACCATCT 1440
Db 1381 CTCGCAAGCATCGGCAAGTCAAGGCTCTCTGTTCTCGGATGATACCCATACCATCT 1440
Qy 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500
Qy 1501 GTCTGGAATAGAAAGGTTACAGAAAAAGTGTGGACAGTGGGTGTCAGAGACTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGTTACAGAAAAAGTGTGGACAGTGGGTGTCAGAGACTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATAAAGAGTTCCTGCTCAAGTCTGTCATACAATTCCTCTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGAGTTCCTGCTCAAGTCTGTCATACAATTCCTCTGGC 1620
Qy 1621 ACATCTTGTGAGAGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680
Db 1621 ACATCTTGTGAGAGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680
Qy 1681 ATCTCATCTATGACAAATTTGGTGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 ATCTCATCTATGACAAATTTGGTGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCTCAGATATAGTACTTCTACCTGTGATTTTTCAGAGATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATATAGTACTTCTACCTGTGATTTTTCAGAGATATGAATATGATCTC 1860
Qy 1861 AAATGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
Db 1861 AAATGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
Qy 1921 GGAAGAGTGAACCGAAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAGAGTGAACCGAAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Qy 1981 GCCGTCAAAATGCTGAAGAGAAAGACACAGCTCTGAAGAGAGGCACTCATGTCAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAGAGAAAGACACAGCTCTGAAGAGAGGCACTCATGTCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAGGCCACAGAAATTTGTGAACCTGCTGGGGCGTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAGGCCACAGAAATTTGTGAACCTGCTGGGGCGTGC 2100

Qy 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGCTATGCTGATCTTCTCAAC 2160
Db 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGCTATGCTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAAGTAAAGAGAAATTTTACAGAGCTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAAGTAAAGAGAAATTTTTCACAGACTTGGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTCAAGTTTACCCCACTTTCCAAATCAGATCCAAATTCAGCATGCTGGTTCAGA 2280
Db 2221 AATTTCAAGTTTACCCCACTTTCCAAATCAGATCCAAATTCAGCATGCTGGTTCAGA 2280
Qy 2281 GAAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTTC 2340
Db 2281 GAAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTTC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTCTTCTTGTGATATCAAGTTGCCAAAGAAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTCTTCTTGTGATATCAAGTTGCCAAAGAAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTTTTCACAGAGACCTGGCCGCCAGGAACGCTGTCACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTTCACAGAGACCTGGCCGCCAGGAACGCTGTCACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640
Qy 2641 GAAAGCATCTACACCATTAAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAATC 2700
Db 2641 GAAAGCATCTACACCATTAAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAATC 2700
Qy 2701 TTCTCACTTGTGTGAATCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760
Db 2701 TTCTCACTTGTGTGAATCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760
Qy 2761 ATTCAAAATGGAATTTAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 ATTCAAAATGGAATTTAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Qy 2821 ATGCAATCTCTGCTGGCTTTTACCTCAAGGAACGCCATCTCCCTCAATTTGACTTCG 2880
Db 2821 ATGCAATCTCTGCTGGCTTTTACCTCAAGGAACGCCATCTCCCTCAATTTGACTTCG 2880
Qy 2881 TTTTGTAGGATGTCAGTGGCAGATGCAAGAAAGCGATGATCAGAAATGTTGATGGCCGT 2940
Db 2881 TTTTGTAGGATGTCAGTGGCAGATGCAAGAAAGCGATGATCAGAAATGTTGATGGCCGT 2940
Qy 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTTCAGCAGAGAGATGGA 3000
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTTCAGCAGAGAGATGGA 3000
Qy 3001 GGGCTACTCTCTCCGAGCTCAGGTGCAAGATTCGTAGAGNACAATTTAGTTTTAAGG 3060
Db 3001 GGGCTACTCTCTCCGAGCTCAGGTGCAAGATTCGTAGAGNACAATTTAGTTTTAAGG 3060
Qy 3061 ACTTCATCCCTCACCTATCTCCCTACAGGCTGTAGATTTACCAAAACAGATTAATTTT 3120
Db 3061 ACTTCATCCCTCACCTATCTCCCTACAGGCTGTAGATTTACCAAAACAGATTAATTTT 3120
Qy 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGAGACTTTTCTTAGAGAGCGTCT 3180
Db 3121 CACTAAAAGAAAATCTATTATCACTGCTTCCAGAGACTTTTCTTAGAGAGCGTCT 3180

QY 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAATCATCTGTCAAGGCAG 3240
 DB |||||||
 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAATCATCTGTCAAGGCAG 3240
 QY 3241 GAGGAGCTGATATGAACTTTATGAGCATTCATCTGCATCAAGCCCTTCTCAGGCCG 3300
 DB |||||||
 3241 GAGGAGCTGATATGAACTTTATGAGCATTCATCTGCATCAAGCCCTTCTCAGGCCG 3300
 QY 3301 GCTTGAGTGAATGTTGCTGACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
 DB |||||||
 3301 GCTTGAGTGAATGTTGCTGACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
 QY 3361 ATTTTCTAAGGAGAACTAATATGATTTTTTAAAGTCTATGTTTAAATATATGTA 3420
 DB |||||||
 3361 ATTTTCTAAGGAGAACTAATATGATTTTTTAAAGTCTATGTTTAAATATATGTA 3420
 QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
 DB |||||||
 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480
 QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
 DB |||||||
 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
 RESULT 6
 AAQ53503
 ID AAQ53503 standard; cDNA; 3501 BP.
 XX
 AC AAQ53503;
 XX
 DT 27-JUN-1994 (first entry)
 XX
 DE Human flk-2 cDNA.
 XX
 KW Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
 KW htk; primitive; totipotent; haematopoietic cell; stem cell;
 KW proliferation; stromal cell; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 58..3039
 FT /*tag= a
 FT /product= flk-2
 FT sig_peptide 58..138
 FT /*tag= b
 FT mat_peptide 139..3036
 FT /*tag= c
 XX
 PN US5270458-A.
 XX
 PD 14-DEC-1993.
 XX
 PF 02-APR-1991; 91US-0679666.
 XX
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 19-NOV-1992; 92US-0977451.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 DR WPI; 1993-405021/50.
 DR P-PSDB; AAR44995.
 XX
 PT Isolated nucleic acid molecules of hematopoietic stem cell
 PT receptor flk-2 - encoding mammalian receptor protein tyrosine

PT kinases expressed in primitive haematopoietic cells
 XX Disclosure; Fig 1b; 60pp; English.
 PS
 CC Nucleic acid sequences coding for murine flk-2 and specified
 CC subfragments of it are claimed. The human flk-2 coding sequence
 CC (i.e. AAQ53503) is also disclosed. The flk-2 polypeptides represent
 CC a new class of receptor protein tyrosine kinases which are
 CC expressed only in primitive haematopoietic cells.
 XX
 SQ Sequence 3501 BP; 1068 A; 708 C; 785 G; 940 T; 0 other;
 Query Match 100.0%; Score 3499.4; DB 14; Length 3501;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAGGCGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGCTCCGGAGGCCATG 60
 DB |||||||
 1 CGAGGCGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGCTCCGGAGGCCATG 60
 QY 61 CCGCGTGGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 DB |||||||
 61 CCGCGTGGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAATCAATCAAG 180
 DB |||||||
 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAATCAATCAAG 180
 QY 181 AACAAATGATTCATCAGTGGGAAATCATCATATCCATCCATGATATCCATCCATG 240
 DB |||||||
 181 AACAAATGATTCATCAGTGGGAAATCATCATATCCATCCATGATATCCATCCATG 240
 QY 241 GACCTCGGGTGTGGTGTGAGACCCCGAGAGCTCAGGAGACAGTACGAGAGCTCCGCTG 300
 DB |||||||
 241 GACCTCGGGTGTGGTGTGAGACCCCGAGAGCTCAGGAGACAGTACGAGAGCTCCGCTG 300
 QY 301 GAAGTGAATGATCTGCTTCCATCAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 360
 DB |||||||
 301 GAAGTGAATGATCTGCTTCCATCAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 360
 QY 361 TCCTGCTCTCTGGGCTTTAAGCACAGCTCCCTGAAATTCGCCAGCCACATTTGATTACAA 420
 DB |||||||
 361 TCCTGCTCTCTGGGCTTTAAGCACAGCTCCCTGAAATTCGCCAGCCACATTTGATTACAA 420
 QY 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAAATTCACAGAACCCCAAGCTGAGAAATAC 480
 DB |||||||
 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAAATTCACAGAACCCCAAGCTGAGAAATAC 480
 QY 481 CTACTTTTATTTCAGAGTGAAGTACCAATTTACAAATATTGTTTACAGTGAATTAAGA 540
 DB |||||||
 481 CTACTTTTATTTCAGAGTGAAGTACCAATTTACAAATATTGTTTACAGTGAATTAAGA 540
 QY 541 AATACCCCTGTTTACACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGAGGCC 600
 DB |||||||
 541 AATACCCCTGTTTACACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGAGGCC 600
 QY 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCTGGAATGGCTGCTTCGATTC 660
 DB |||||||
 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCTGGAATGGCTGCTTCGATTC 660
 QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCCT 720
 DB |||||||
 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCCT 720
 QY 721 CATGAATTTTGGGAGCGACATAGGCTGTGCCAGAAATGAATGAGCGAGGAAATGC 780
 DB |||||||
 721 CATGAATTTTGGGAGCGACATAGGCTGTGCCAGAAATGAATGAGCGAGGAAATGC 780
 QY 781 ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCCTCAGACCATTTGCCACAAATTT 840
 DB |||||||
 781 ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCCTCAGACCATTTGCCACAAATTT 840
 QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900

Db 841 CTTAAAGTAGGGAAACCTTATGATAGGTGCAAGCTGTTCATGTGAAACCATGGATTG 900
Qy 901 GGGCTCACCTGGGAATTTAGAAACAAGACACTCAGAGGGCAACTACTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTTAGAAACAAGACACTCAGAGGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAACAACACTATGATACGGATTCTGTGCTTTGTTGTTATCATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAACAAGACTATGATACGGATTCTGTGCTTTGTTGTTATCATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCTGTAGAAAGGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAAT 1140
Db 1081 GTTACCATCTGTAGAAAGGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAAT 1140
Qy 1141 GACCAATATGAAGATTTTCTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
Db 1141 GACCAATATGAAGATTTTCTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
Qy 1201 ACSTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTGTGATAACGGATAC 1260
Db 1201 ACSTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTGTGATAACGGATAC 1260
Qy 1261 AGCATATCCAAGTTTTCGAATCATATAAGCACAGCCAGGAGATATATATCCATGCGAA 1320
Db 1261 AGCATATCCAAGTTTTCGAATCATATAAGCACAGCCAGGAGATATATATCCATGCGAA 1320
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGTAATATAGAAGGAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGTAATATAGAAGGAACCTCAAGTG 1380
Qy 1381 CTGCGAAGCATCGGCAAGTCTGAGGCTCTGTTCTCGGATGATACCCATTTACCATCT 1440
Db 1381 CTGCGAAGCATCGGCAAGTCTGAGGCTCTGTTCTCGGATGATACCCATTTACCATCT 1440
Qy 1441 TGGACCTTGGAAAGAGTCTCAGACAAGTCTCCCACTGCAAGAGAGATCACAGAAGGA 1500
Db 1441 TGGACCTTGGAAAGAGTCTCAGACAAGTCTCCCACTGCAAGAGAGATCACAGAAGGA 1500
Qy 1501 GTCTGGAATAGAAAGGCTACAGAAAAGTGTTCGACAGTGGGTCTCGACGAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGCTACAGAAAAGTGTTCGACAGTGGGTCTCGACGAGTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAATTCCTCTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGCATACAATTCCTCTGGC 1620
Qy 1621 ACATCTTGTGAGAGTATCCCTTTAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680
Db 1621 ACATCTTGTGAGAGTATCCCTTTAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680
Qy 1681 ATCTCATCTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 ATCTCATCTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 ATTTGTACAGATACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTACAGATACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCTCAGATAATGAGTACTCTACGTTGATTTACAGAGATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAATGAGTACTCTACGTTGATTTACAGAGATATGAATATGATCTC 1860
Qy 1861 AAATGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGSTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGSTACTAGGATCAGGTGCTTTT 1920
Qy 1921 GGAAGTGTATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
Db 1921 GGAAGTGTATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980

Db 1921 GGAAGTGTATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
Qy 1981 GCCGTCAAAATGCTGAAAGAAAAGCACACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAAGAAAAGCACACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Qy 2041 CTCGAAGTATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
Db 2041 CTCGAAGTATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
Qy 2101 ACACGTCTCAGGACCAATTTACTTTGATTTGTAATCTGCTGCTGATCTTCTCAAC 2160
Db 2101 ACACGTCTCAGGACCAATTTACTTTGATTTGTAATCTGCTGCTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTTCAGTTTACCCTTTCCTCAATTCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
Db 2221 AATTTTCAGTTTACCCTTTCCTCAATTCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
Qy 2281 GAACTTCAGATACACCCGACTTCGATCAAAATCTCAGGCTTTTATGGGAATTCATTTAC 2340
Db 2281 GAACTTCAGATACACCCGACTTCGATCAAAATCTCAGGCTTTTATGGGAATTCATTTAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAAGAGAGGAGACTTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAAGAGAGGAGACTTTGAAT 2400
Qy 2401 GTGCTTACATTTGAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTTGGCCGAGGAAGTGTCTGTCACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTTGGCCGAGGAAGTGTCTGTCACCCAC 2520
Qy 2521 GGGAAAGTGTGAAGATATGACATTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580
Db 2521 GGGAAAGTGTGAAGATATGACATTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580
Qy 2581 TATGTTGTCAGGGCAATGCCGCTGCTGTAAATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 TATGTTGTCAGGGCAATGCCGCTGCTGTAAATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 GAAGCATCTACACCATTTAAGATGATGCTGCTCATATGGAATATTTACTGTGGGAATC 2700
Db 2641 GAAGCATCTACACCATTTAAGATGATGCTGCTCATATGGAATATTTACTGTGGGAATC 2700
Qy 2701 TTCTCAGTTGGTGAATCTTACCTTGGCATTTCCGGTTGATGCTAACTTCTTACAACTG 2760
Db 2701 TTCTCAGTTGGTGAATCTTACCTTGGCATTTCCGGTTGATGCTAACTTCTTACAACTG 2760
Qy 2761 ATTCAAAATGATTTAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 ATTCAAAATGATTTAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Qy 2821 ATGCAATCTCTGGGCTTTTGAATCAAGGAAGCCATCTTCCCTAATTTGATCTCG 2880
Db 2821 ATGCAATCTCTGGGCTTTTGAATCAAGGAAGCCATCTTCCCTAATTTGATCTCG 2880
Qy 2881 TTTTGTAGATGTCAGCTGGCAGATGCAGAAGAGGATGATGATGATGATGATGATGATGATGAT 2940
Db 2881 TTTTGTAGATGTCAGCTGGCAGATGCAGAAGAGGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 GTTTCGGAATGTCCTCAGCTTACCAACAGGAGGCTTTTTCAGAGAGAGATGATGATGATGATGAT 3000
Db 2941 GTTTCGGAATGTCCTCAGCTTACCAACAGGAGGCTTTTTCAGAGAGAGATGATGATGATGATGAT 3000
Qy 3001 GGGTACTCTCTCCGAGGCTCAGTCTGGAAGTTTCTGAGAGGAACAAATTTAGTTTAAAGG 3060
Db 3001 GGGTACTCTCTCCGAGGCTCAGTCTGGAAGTTTCTGAGAGGAACAAATTTAGTTTAAAGG 3060

Qy	3061	ACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Qy	3121	CACTTAAAGAAAATCTATTATCAACTGCTGCTTCCACAGAGCTTTTCTCTAGAAGCCGCTCT	3180
Db	3121	CACTTAAAGAAAATCTATTATCAACTGCTGCTTCCACAGAGCTTTTCTCTAGAAGCCGCTCT	3180
Qy	3181	GGCTTTTACTCTTGTTTTCAAAGGAGCTTTTGTGTAATCAATCATCTGTCACAAGGCAG	3240
Db	3181	GGCTTTTACTCTTGTTTTCAAAGGAGCTTTTGTGTAATCAATCATCTGTCACAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCATTGAGTGAATTTGTACCTCGAAGTACATATATCTCTGAATACATAAACAANAAGC	3360
Db	3301	GCATTGAGTGAATTTGTACCTCGAAGTACATATATCTCTGAATACATAAACAANAAGC	3360
Qy	3361	ATTTTGCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTAAATAATAATGTTAAA	3420
Db	3361	ATTTTGCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTAAATAATAATGTTAAA	3420
Qy	3421	TTTTTTCAGCTATTTAGTGATATATTTTATCGGTGGGAATAAATTTCTCTACTACAGAAAAA	3480
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTATCGGTGGGAATAAATTTCTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 7

AAQ40915	standard; cDNA; 3501 BP.
AAQ40915	
AAQ40915	
19-OCT-1993	(first entry)
Human flk-2 cDNA.	
Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive; hematopoietic cell; mature; family; conserved; region;	
catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;	
thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony; hierarchy; transduction; T-lymphoid; lineage; ss.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	58..3039
	/*tag= a
sig_peptide	58..138
	/*tag= b
mat_peptide	/note= "Hydrophobic leader sequence"
	139..3036
	/*tag= C
WO9310136-A.	
27-MAY-1993.	
16-NOV-1992;	92WO-US09893.
15-NOV-1991;	91US-0793065.
(UYPR-) UNIV PRINCETON.	
Lemischka IR;	

Db 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGTGCTTTGCGATTCA 660
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGCTGTTTAAAGAGGAGGAAAAGTCTT 720
Db 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGCTGTTTAAAGAGGAGGAAAAGTCTT 720
Qy 721 CATGAATTAATTTGGAGCGGACATAAGGTGTGTCGCCAGAAATGAACCTGGGCGAGGAATGC 780
Db 721 CATGAATTAATTTGGAGCGGACATAAGGTGTGTCGCCAGAAATGAACCTGGGCGAGGAATGC 780
Qy 781 ACCAGGCTGTTCAACATAGATCTAAATCAAACTCCTCAGACCAATGTCGCCACAATATTT 840
Db 781 ACCAGGCTGTTCAACATAGATCTAAATCAAACTCCTCAGACCAATGTCGCCACAATATTT 840
Qy 841 CTTAAAGTAGGGNACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGGATTTC 900
Db 841 CTTAAAGTAGGGNACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGGATTTC 900
Qy 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTATCATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTATCATCATCAGTGGCA 1020
Qy 1021 AGAACCACACCGGATACTACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAACCACACCGGATACTACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCTGAGGAAAGGATTTATAATGCTACCAATTCACGTGAAGATTATGAATTT 1140
Db 1081 GTTACCATCTGAGGAAAGGATTTATAATGCTACCAATTCACGTGAAGATTATGAATTT 1140
Qy 1141 GACCAATPATGAAGAGTTTGTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200
Db 1141 GACCAATPATGAAGAGTTTGTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200
Qy 1201 ACGTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGATTAACGGATAC 1260
Db 1201 ACGTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGATTAACGGATAC 1260
Qy 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACAGGAGGATATATATCCATGCGAGAA 1320
Db 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACAGGAGGATATATATCCATGCGAGAA 1320
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAAATATAAGAAGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAAATATAAGAAGAAACCTCAAGTG 1380
Qy 1381 CTGCGAAGCATCGGCAAGTCAAGGCTCGTGTTCGCGATGATACCCCATTTACCATCT 1440
Db 1381 CTGCGAAGCATCGGCAAGTCAAGGCTCGTGTTCGCGATGATACCCCATTTACCATCT 1440
Qy 1441 TGGACCTGGAGAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500
Db 1441 TGGACCTGGAGAGTGTTCAGACAAGTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500
Qy 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTGTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTGTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 AACATGATGAAGCCATAAAGGTTCTTGGTCAAGTGTCTGTGATACAAATTCCTTGGC 1620
Db 1561 AACATGATGAAGCCATAAAGGTTCTTGGTCAAGTGTCTGTGATACAAATTCCTTGGC 1620
Qy 1621 ACATCTTGTGAGAGCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAC 1680
Db 1621 ACATCTTGTGAGAGCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAC 1680
Qy 1681 ATCTCATTTCTATGCAAAATTTGGTGTGTCCTCTCTTCTATTGTCGTTTAAACCTCTGTA 1740
Db 1681 ATCTCATTTCTATGCAAAATTTGGTGTGTCCTCTCTTCTATTGTCGTTTAAACCTCTGTA 1740

Qy 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTTGATTTTTCAGAGAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTTGATTTTTCAGAGAATATGAATATGATCTC 1860
Qy 1861 AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
Db 1861 AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
Qy 1921 GGAAGAGTGTGAACGCAACAGCTTATGSAATTAGCAAAACAGGAGTCTCAATCAGGTT 1980
Db 1921 GGAAGAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCAGGTT 1980
Qy 1981 GCCGTCAAAATGCTCAAGAAAAAGCAGACAGCTCTGAAAAGAGAGGCACTCATGTCAAGAA 2040
Db 1981 GCCGTCAAAATGCTCAAGAAAAAGCAGACAGCTCTGAAAAGAGAGGCACTCATGTCAAGAA 2040
Qy 2041 CTCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCTG 2100
Db 2041 CTCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCTG 2100
Qy 2101 ACCTGTGAGCAACCAATTTACTTGTATTTTGAATACTGTTGCTATGGTGATCTTCAAC 2160
Db 2101 ACCTGTGAGCAACCAATTTACTTGTATTTTGAATACTGTTGCTATGGTGATCTTCAAC 2160
Qy 2161 TATCTTAAGAAATTAAGAGAAAAATTTTTCACAGGACTTGGACAGAGATTTTCAAGAACAC 2220
Db 2161 TATCTAAGAAATTAAGAGAAAAATTTTTCACAGGACTTGGACAGAGATTTTCAAGAACAC 2220
Qy 2221 AATTTAGTTTTTACCCTTCCAAATCACATCCAAATTCAGCATGCTGCTGTTCAAGA 2280
Db 2221 AATTTAGTTTTTACCCTTCCAAATCACATCCAAATTCAGCATGCTGCTGTTCAAGA 2280
Qy 2281 GAAGTTGATGATACACCCGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTTAC 2340
Db 2281 GAAGTTGATGATACACCCGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTTAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTTGGAAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTTGGAAAGAGAGGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTTTCAGAGACCTGGCCGCGCAGGAAGCTGCTTGTCAACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTCAGAGACCTGGCCGCGCAGGAAGCTGCTTGTCAACCCAC 2520
Qy 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACAC 2580
Db 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCCCGGAAAGGCTTGT 2640
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCCCGGAAAGGCTTGT 2640
Qy 2641 GAAGGATCTACACCATTAAGAGTATGCTGGTTCATATGGGAATTTACTGTGGGAATC 2700
Db 2641 GAAGGATCTACACCATTAAGAGTATGCTGGTTCATATGGGAATTTACTGTGGGAATC 2700
Qy 2701 TTCTCAGTTGTTGAATCCTTACCTGGCATTCGGTTTGAAGTCTTCAACTTCTACAACTG 2760
Db 2701 TTCTCAGTTGTTGAATCCTTACCTGGCATTCGGTTTGAAGTCTTCAACTTCTACAACTG 2760
Qy 2761 ATTCAAAATGGAATTTAAAATGGATCAGGCAATTTTATGCTACAGAGAATAATACATATA 2820
Db 2761 ATTCAAAATGGAATTTAAAATGGATCAGGCAATTTTATGCTACAGAGAATAATACATATA 2820

QY	2761	ATTCAAATGGATTTAAAATGGATCAGCCATTTATGCTACAGAAGAATATACATTATA	2820
DB	2761		2820
QY	2821	ATGCAATCCCTGGGCTTTTGACTCAAGGAACGCCATCTTCCCTAAATTTGACTTCG	2880
DB	2821		2880
QY	2881	TTTTTAGGATGTCAGCTGGCAGATGCAGAAGAAGCGATATCAGAATGTTGGATGCCGT	2940
DB	2881		2940
QY	2941	GTTTCGGAATGTCCTCACACCTACCAAAACAGCGACCTTTCAGCAGAGAGATGGATTG	3000
DB	2941		3000
QY	3001	GGGCTACTCTCGCGAGGCTCAGGTCGAAGATTCGTAGAGNACAATTTAGCTTTTAAGG	3060
DB	3001		3060
QY	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTAACAAAAACAAGATTAAATTCAT	3120
DB	3061		3120
QY	3121	CACATAAAGAAAAATCTATTATCAACTGCTGCTCCACAGACTTTTCTAGAGCGCGTCT	3180
DB	3121		3180
QY	3181	CGGTTTACTCTGTTTTCAAGGGACCTTTTGTAAATCAAATCATCTGTCACAAGGCAG	3240
DB	3181		3240
QY	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAAGGCCCTTCAGGCCCG	3300
DB	3241		3300
QY	3301	GCTTGAGTGAATGTGTACCTGGAAGTACAGTATATCTTGTGAATACATACATAAAACAAGC	3360
DB	3301		3360
QY	3361	ATTTTGCTTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTTAAATAAATATGTAAA	3420
DB	3361		3420
QY	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
DB	3421		3480
QY	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
DB	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 9
AAQ91536
ID AAQ91536 standard: DNA: 3476 BP.

XX
DT 29-DEC-1995 (first entry)

XX DE Human STK-1 cDNA.

XX. STK-1; receptor PTK; protein tyrosine kinase; ss.
KW

OS Homo sapiens.

XX	Key	Location/Qualifiers
FH		

ET	58..2976
CDS	.

FT 1000-1000

```
FT misc_feature 1..5/
FT /*tag=
```

```
FT /label= cl
```

FT	misc_feature	121..148	
FT		/tag= C	
FT		/label= see above	
XX			
XX	W09519175-A.		
XX			
PD	20-JUL-1995.		
XX			
XX	06-JAN-1995;	95WO-US00176.	
XX			
PR	14-JAN-1994;	94US-0183211.	
XX			
XX	(UYJO) UNIV JOHNS HOPKINS.		
PA	(UYPE-) UNIV PENNSYLVANIA.		
XX			
PI	Civin Cl, Gewirtz AM, Small D;		
XX			
DR	WPI; 1995-263709/34.		
DR	P-PSDB; AAR75961.		
XX			
PT	Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide		
PT	used to treat neoplastic diseases and as bone marrow purging agents for		
PT	treating leukaemia and neoplasma		
XX			
PS	Claim 8; Fig 1; 66pp; English.		
XX			
CC	The STK-1 gene encodes a receptor PTK which is expressed in		
CC	proliferating hematopoietic stem cells but not in quiescent stem		
CC	cells. The STK-1 gene is also expressed in certain malignant cells		
CC	of non-hematopoietic origin. An antisense oligo specific for STK-1		
CC	is an oligo having a sequence (i) capable of forming a stable		
CC	triplex with a portion of the STK-1 gene, or (ii) capable of forming		
CC	a stable duplex with a portion of an mRNA transcript of the STK-1		
CC	gene. Antisense oligos capable of forming a stable duplex with a		
CC	portion of a STK-1 mRNA transcript are given in AA091536 FT and in		
CC	AA091537 and AA091538. The antisense oligos of the invention are useful		
CC	in the treatment of hematologic malignancies characterised by		
CC	STK-1 expression.		
XX			
SO	Sequence 3476 BP: 1044 A; 709 C; 783 G; 940 T; 0 other;		

Query Match	99.2%;	Score 3474.4;	DB 16;	Length 3476;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3475;	Conservative	0; Mismatches 1;	Indels 0;	Gaps 0;
QY	1	CGAGGCGCATCCGAGGCGTGGCGCGCGCCCTCTGGGGACCCCGGCTCCGGAGGCCATG	60	
Db	1			
QY	61	CGAGGCGCATCCGAGGCGTGGCGCGCGCCCTCTGGGGACCCCGGCTCCGGAGGCCATG	60	
Db	61			
QY	121	CGGCGCTTGGCGCGACCGCGGCGCATCGTGGTGTCTGCTGTTGTTTTTCTGCAATGATA	120	
Db	121			
QY	121	TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTTTAAATCAATCATAAAG	180	
Db	121			
QY	181	AACAATGATTATCATAGTGGGGAAGTCATCATATFCCCATGGTATCAGATATCCCGGAA	240	
Db	181			
QY	241	GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGSACAGTGTACGAAGCTGCCGCTGTG	300	
Db	241			
QY	301	GAAGTGGATGATCTGCTTCCATCACATGCAAGTGTGGTGCATGCCCCAGGGGAACATT	360	
Db	301			
QY	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGCAGCCACATTTTGATTTACAA	420	
Db	361			
QY	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGCAGCCACATTTTGATTTACAA	420	
Db	361			

QY 421 AACAGAGAGTGTTCATCGGTGCTATTTTGAATGACAGAAACCCAGCTGGAGAATAC 480
DB 421 AACAGAGAGTGTTCATCGGTGCTATTTTGAATGACAGAAACCCAGCTGGAGAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTGTTTACAGTGAATAGA 540
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTGTTTACAGTGAATAGA 540
QY 541 AATACCCCTGCTTTACACATTAAGAGACCTTTACTTTAGAAAAATGAAAAACCCAGGACCC 600
DB 541 AATACCCCTGCTTTACACATTAAGAGACCTTTACTTTAGAAAAATGAAAAACCCAGGACCC 600
QY 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGTCTTTGCGATTCA 660
DB 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGTCTTTGCGATTCA 660
QY 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAAAGTGCTT 720
DB 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAAAGTGCTT 720
QY 721 CATGAATTTATTTGGACGGACATAAGGTGCTGTGCCAGAAATGAACCTGGGAGGGAATGC 780
DB 721 CATGAATTTATTTGGACGGACATAAGGTGCTGTGCCAGAAATGAACCTGGGAGGGAATGC 780
QY 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCTCAGACACATTCGCCACAATATT 840
DB 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCTCAGACACATTCGCCACAATATT 840
QY 841 CTTAAAGTAGGGGAACCTTTATGATAGGTGCAAGCTGTTCATGTGAACCATGGATT 900
DB 841 CTTAAAGTAGGGGAACCTTTATGATAGGTGCAAGCTGTTCATGTGAACCATGGATT 900
QY 901 GGGCTCACCTGGGAATTAGAAAAACAAAGCACTCAGAGGAGGCAACTACTTTGAGATGAGT 960
DB 901 GGGCTCACCTGGGAATTAGAAAAACAAAGCACTCAGAGGAGGCAACTACTTTGAGATGAGT 960
QY 961 ACCTATTCAACAAACAGAACTATGATCGGATCTGTTTGTGTTATCATCAGTGCA 1020
DB 961 ACCTATTCAACAAACAGAACTATGATCGGATCTGTTTGTGTTATCATCAGTGCA 1020
QY 1021 AGAACGACACCGGATCTACACATGTTCCCTTCAAGAGTCCAGTCAATCAGCTTG 1080
DB 1021 AGAACGACACCGGATCTACACATGTTCCCTTCAAGAGTCCAGTCAATCAGCTTG 1080
QY 1081 GTTACCATCTGAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
DB 1081 GTTACCATCTGAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
QY 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200
DB 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200
QY 1201 ACGTGGACCTTCTCGAANAATCATTTCCCTTGTGAGCAAAAGGCTTGTATACGGATAC 1260
DB 1201 ACGTGGACCTTCTCTCGAANAATCATTTCCCTTGTGAGCAAAAGGCTTGTATACGGATAC 1260
QY 1261 AGCATATCCAAAGTTTGTCAATCAAGCAGCAGGAGGAATATATATTCATCGACAAA 1320
DB 1261 AGCATATCCAAAGTTTGTCAATCAAGCAGCAGGAGGAATATATATTCATCGACAAA 1320
QY 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAAATATAAGAAAGAAACCTCAAGTG 1380
DB 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAAATATAAGAAAGAAACCTCAAGTG 1380
QY 1381 CTCGAGAAGCATCGGCAAGTACAGCGCTCTGTTCTCGGATGGATACCCATTTACCATCT 1440
DB 1381 CTCGAGAAGCATCGGCAAGTACAGCGCTCTGTTCTCGGATGGATACCCATTTACCATCT 1440
QY 1441 TGGACCTGGAGAGAGTGTTCAGACAAGTCTCCACCTCAGCAGAGAGATCACAGAAGA 1500
DB 1441 TGGACCTGGAGAGAGTGTTCAGACAAGTCTCCACCTCAGCAGAGAGATCACAGAAGA 1500
QY 1501 GTCGGAATAGAAAGGCTTAACAGAAAAAGTGTGACAGAGTGGGTGTCGAGCAGTACTCTA 1560

DB 1501 GTCTCGAATAGAAAGGCTTAACAGAAAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGTCATACAAATTCCTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGTCATACAAATTCCTTGGC 1620
QY 1621 ACATCTTGTGAGACGATCCTTTTAAACTCTCAGGCCCCCTCCCTTTTCAAGACAAC 1680
DB 1621 ACATCTTGTGAGACGATCCTTTTAAACTCTCAGGCCCCCTCCCTTTTCAAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 1741 ATTTGTCAACAGTACAAAAGCAATTTAGTGTGAAGCCAGCTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTCAACAGTACAAAAGCAATTTAGTGTGAAGCCAGCTACAGATGGTACAGGTG 1800
QY 1801 ACCGCTCTCTCAGATAATAGTACTTCTACGTTGATTTGAGAGAATATGAATATGATCTC 1860
DB 1801 ACCGCTCTCTCAGATAATAGTACTTCTACGTTGATTTGAGAGAATATGAATATGATCTC 1860
QY 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAAAGTACTAGGATCAGGTGCTTT 1920
DB 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAAAGTACTAGGATCAGGTGCTTT 1920
QY 1921 GGAAAGTGTGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
DB 1921 GGAAAGTGTGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
QY 2041 CTCAGATGATGACCCAGCTGGAAAGCCACAGAAATATTGTTGAACTGCTGGGGGGTGC 2100
DB 2041 CTCAGATGATGACCCAGCTGGAAAGCCACAGAAATATTGTTGAACTGCTGGGGGGTGC 2100
QY 2101 ACACGTGTCAGGACCAATTTACTTGTGTTTGAATCTGTTGCTATGCTGCTCTCTCAAC 2160
DB 2101 ACACGTGTCAGGACCAATTTACTTGTGTTTGAATCTGTTGCTATGCTGCTCTCTCAAC 2160
QY 2161 TATCTAAGAAGTAAAGAGAAAAATTTTCAAGACCTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAAGTAAAGAGAAAAATTTTCAAGACCTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTCAAGTTTACCCACCTTTTCCAAATCACAATCCAAATTCAGCATGCTGTTTCAAGA 2280
DB 2221 AATTTCAAGTTTACCCACCTTTTCCAAATCACAATCCAAATTCAGCATGCTGTTTCAAGA 2280
QY 2281 GAAGTTCAGATACCCCGGACTCGGATCAAAATCTCAGGCTTCATCGGAAATTCATTTTCA 2340
DB 2281 GAAGTTCAGATACCCCGGACTCGGATCAAAATCTCAGGCTTCATCGGAAATTCATTTTCA 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCCAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCCAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTTGGATATCAAGTTGCCAAAGAAATTTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTTGGATATCAAGTTGCCAAAGAAATTTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTTGGCGCCAGGAACTGCTGTCACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTTGGCGCCAGGAACTGCTGTCACCCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGACATTTGGATTGGCTGCGAGATATCATGATGATTTCAAC 2580
DB 2521 GGGAAAGTGGTGAAGATATGACATTTGGATTGGCTGCGAGATATCATGATGATTTCAAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCCTGCTGCTGTTAAATGATGGCCCCCAGAAAGCTGTTT 2640

XX	Key	Location/Qualifiers
PH	CDS	58..3039
FT		/*tag= a
FT		/product= flk2/flt3_tyrosine_kinase_receptor
XX		
PN		W09527062-A1.
XX		
PD		12-OCT-1995.
XX		
PF		23-MAR-1995; 95WO-US03718.
XX		
PR		04-APR-1994; 94US-0222299.
XX		
PA		(GETH) GENENTECH INC.
XX		
PI		Bennett BD, Broz SD, Matthews W, Zeigler FC;
XX		
DR		WPI; 1995-358636/46.
XX		P-PSDB; AAR81869.
PT		Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
PT		- enhances proliferation of haematopoietic stem cells, in the
PT		treatment of hypoplasia, anaemia, etc.
XX		
PS		Disclosure; Page 41-44; 59pp; English.
XX		
CC		DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2
CC		(also called flt3) (AAR81868) was obtd. by RT-PCR amplification of RNA
CC		isolated from mid-gestation mouse foetal livers using primers based
CC		on the murine flt3 sequence, and subcloning of the product into pRK5.1.
CC		An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was
CC		constructed and used to raise agonist antibodies able to bind to, and
CC		activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and
CC		amino acid sequence (AAR81869) are also given.
XX		
SQ		Sequence 3475 BP; 1042 A; 709 C; 784 G; 940 T; 0 other;
Query Match 99.1%; Score 3470.2; DB 16; Length 3475;		
Best Local Similarity 99.9%; Pred. No. 0;		
Matches 3472; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
QY	1	CGAGGCGCATCCGAGGGCTGGCCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG 60
Db	1	CGAGGCGCATCCGAGGGCTGGCCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG 60
QY	61	CCGGCGTTGGCGCGGACCGCGGCGCCCTGCTGCTGCTTTTCTGCAATGATA 120
Db	61	CCGGCGTTGGCGCGGACCGCGGCGCCCTGCTGCTGCTTTTCTGCAATGATA 120
QY	121	TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATG 180
Db	121	TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATG 180
QY	181	AACAATGATTTCATCAGTGGGAAATCATCATCATATCCCATGGTATCAGAAATCCCGGAA 240
Db	181	AACAATGATTTCATCAGTGGGAAATCATCATCATATCCCATGGTATCAGAAATCCCGGAA 240
QY	241	GACCTCGGTGTCGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTCCCGCTGTG 300
Db	241	GACCTCGGTGTCGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTCCCGCTGTG 300
QY	301	GAAGTGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGAAACATT 360
Db	301	GAAGTGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGAAACATT 360
QY	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGACGCCACATTTTGATTACAA 420
Db	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGACGCCACATTTTGATTACAA 420
QY	421	AACAGAGGAGTGTGTTTCCATGGTCAATTTGAAAATGACAGAAACCCAGCTGGAGATAC 480
Db	421	AACAGAGGAGTGTGTTTCCATGGTCAATTTGAAAATGACAGAAACCCAGCTGGAGATAC 480

Db	601	 TCACGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTG	660
Qy	718	CTTTCATGAATTATTTGGGACGGACATAAAGTGTCTGCCAGAAATGAATCACTGGCGACGGAA	777
Db	661	CTTTCATGAATTATTTGGGACGGACATAAAGTGTCTGCCAGAAATGAATCACTGGCGACGGAA	720
Qy	778	TGCACGAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTA	837
Db	721	TGCACGAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTA	780
Qy	838	TTTCTTTAAAGTAGGGNAACCTTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGA	897
Db	781	TTTCTTTAAAGTAGGGNAACCTTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGA	840
Qy	898	TTCCGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTTGAGATG	957
Db	841	TTCCGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTTGAGATG	900
Qy	958	AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTCTTTGTATCATCAGTG	1017
Db	901	AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTCTTTGTATCATCAGTG	960
Qy	1018	GCAAGAAACGACACCGGATCTACACTTGTTCCTCTTCAAAAGCATCCAGTCCATCAGCT	1077
Db	961	GCAAGAAACGACACCGGATCTACACTTGTTCCTCTTCAAAAGCATCCAGTCCATCAGCT	1020
Qy	1078	TTTGTTTACCATCGTAGGAAGGGATTATAAATGCTTACCAATTCAGTCAAGATTATGAA	1137
Db	1021	TTTGTTTACCATCGTAGGAAGGGATTATAAATGCTTACCAATTCAGTCAAGATTATGAA	1080
Qy	1138	ATTGACCAATATGAAGAGTTTCTGTTTCTCTGTCAGGTTTAAAGCCTACCCACAAATCAGA	1197
Db	1081	ATTGACCAATATGAAGAGTTTCTGTTTCTCTGTCAGGTTTAAAGCCTACCCACAAATCAGA	1140
Qy	1198	TGTACGTGGACCTTCTCTCGAAATCATTTCCCTTTGTGAGCAAAAGGCTTGTGATAACGGA	1257
Db	1141	TGTACGTGGACCTTCTCTCGAAATCATTTCCCTTGTGAGCAAAAGGCTTGTGATAACGGA	1200
Qy	1258	TACAGCATATCCAAGTTTGTCAATCATAAAGCACCGCAGGAGAAATATATTTCCATGCA	1317
Db	1201	TACAGCATATCCAAGTTTGTCAATCATAAAGCACCGCAGGAGAAATATATTTCCATGCA	1260
Qy	1318	GAATAATGATGATGCCCAATTTACCAAAATGTTCCACGCTGAATATAAGAGAAACCTCAA	1377
Db	1261	GAATAATGATGATGCCCAATTTACCAAAATGTTCCACGCTGTATATAGAGAAACCTCAA	1320
Qy	1378	GTGCTCGCAGAGCATCGGCAATCGAGCGTCTGTTTCTCGGATGGATACCCATTACCA	1437
Db	1321	GTGCTCGCAGAGCATCGGCAATCGAGCGTCTGTTTCTCGGATGGATACCCATTACCA	1380
Qy	1438	TCCTTGGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCACACAGAGATACAGAA	1497
Db	1381	TCCTTGGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCACACAGAGATACAGAA	1440
Qy	1498	GGAGTCTGGAATAGAAGGCTTAACAGAAAAGTGTTTGGACAGTGGGTGTCGAGCAGTACT	1557
Db	1441	GGAGTCTGGAATAGAAGGCTTAACAGAAAAGTGTTTGGACAGTGGGTGTCGAGCAGTACT	1500
Qy	1558	CTAAACATGAGTGAAGCCATAAAAGGTTCCCTGGTCAAGTGTGTGCATACAAATTCCTTT	1617
Db	1501	CTAAACATGAGTGAAGCCATAAAAGGTTCCCTGGTCAAGTGTGTGCATACAAATTCCTTT	1560
Qy	1618	GGCACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCTCAAGAC	1677
Db	1561	GGCACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCTCAAGAC	1620
Qy	1678	AACATCTCATCTTATGCAACAAATGGTGTGTTGTCTCTCTCTTCATGTGCTGTTTTAACCCCTG	1737
Db	1621	AACATCTCATCTTATGCAACAAATGGTGTGTTGTCTCTCTCTTCATGTGCTGTTTTAACCCCTG	1680
Qy	1738	CTAATTTGTGCACAAAGTACAAAAACAAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797

Db	1681	CTAATTTGTCACAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
Qy	1798	GTGACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTACAGAGAATATGAATATGAT	1857
Db	1741	GTGACCGGATCCTCAGATAATGAGTACTTCTACGTTGATTTACAGAGAATATGAATATGAT	1800
Qy	1858	CTCAATGGGAGTTTCCACAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT	1917
Db	1801	CTCAATGGGAGTTTCCACAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT	1860
Qy	1918	TTTGGAAAAAGTGATGAACGCAACAGCTTATGGAAATTAGCAAAACAGAGGTCTCAATCCAG	1977
Db	1861	TTTGGAAAAAGTGATGAACGCAACAGCTTATGGAAATTAGCAAAACAGAGGTCTCAATCCAG	1920
Qy	1978	GTTCCGCTCAAAATGCTGAAAGAAAAACAGACAGCTCTCAAGAGAGGACATCATGTCA	2037
Db	1921	GTTACCCTCAAAATGCTGAAAGAAAAACAGACAGCTCTCAAGAGAGGACATCATGTCA	1980
Qy	2038	GAATCTAAGATGATGCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGGGG	2099
Db	1981	GAATCTAAGATGATGCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGGGG	2040
Qy	2098	TGCACACTGTCAAGGACCAATTTACTTCGATTTTGAATACTTGTGCTATGGTGATCTTCTC	2157
Db	2041	TGCACACTGTCAAGGACCAATTTACTTCGATTTTGAATACTTGTGCTATGGTGATCTTCTC	2100
Qy	2158	AACATCTAAGAAGTAAACAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA	2217
Db	2101	AACATCTAAGAAGTAAACAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAAATTTACGTTTTTACCCCACTTTTCCAAATCACATCCAAATTCAGCATGCTTGTTCA	2277
Db	2161	CACAAATTTACGTTTTTACCCCACTTTTCCAAATCACATCCAAATTCAGCATGCTTGTTCA	2220
Qy	2278	AGAGAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTTCA	2337
Db	2221	AGAGAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTTCA	2280
Qy	2338	CACCTCAAGATGAAATTTGAATATGAAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTG	2397
Db	2281	CACCTCAAGATGAAATTTGAATATGAAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTG	2340
Qy	2398	AATGTGCTTACATTTGAAGATCTCTTTGCTTTGCATATCAAGTGTGCCAAGGAATGGAA	2457
Db	2341	AATGTGCTTACATTTGAAGATCTCTTTGCTTTGCATATCAAGTGTGCCAAGGAATGGAA	2400
Qy	2458	TTTCTGGAATTTAAGTCGTGTGTTCACAGAGACTTGGCCGCCAGGAACGTGCTTGTCAAC	2517
Db	2401	TTTCTGGAATTTAAGTCGTGTGTTCACAGAGACTTGGCCGCCAGGAACGTGCTTGTCAAC	2460
Qy	2518	CACGGGAAGTGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGACTGATTCC	2577
Db	2461	CACGGGAAGTGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGACTGATTCC	2520
Qy	2578	AACATATGTTGTACGGGGCAATGCCGCTGCCTCTGCTTAAATGGATGGCCCCCGCAAGGCTG	2637
Db	2521	AACATATGTTGTACGGGGCAATGCCGCTGCCTCTGCTTAAATGGATGGCCCCCGCAAGGCTG	2580
Qy	2638	TTTGAAGGCATCTACACCAATTAAGATGTGTCTGTGCATATGGAATATTACTGTGGGAA	2697
Db	2581	TTTGAAGGCATCTACACCAATTAAGATGTGTCTGTGCATATGGAATATTACTGTGGGAA	2640
Qy	2698	ATCTCTCACTTGGTGTGAATCTTACCTCCGGAATCCCGTTGATGCTCACTCACTTCTACAAA	2757
Db	2641	ATCTCTCACTTGGTGTGAATCTTACCTCCGGAATCCCGTTGATGCTCACTCACTTCTACAAA	2700
Qy	2758	CTGATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTCACAGAAGAAATATACATT	2817
Db	2701	CTGATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTCACAGAAGAAATATACATT	2760
Qy	2818	ATAATGAATCCTGCTGGGCTTTTGAATCAAGGAAAGCGGCATCTTCCCTAAATTTGACT	2877
Db	2761	ATAATGAATCCTGCTGGGCTTTTGAATCAAGGAAAGCGGCATCTTCCCTAAATTTGACT	2820

```
QY 2878 TCCTTTTAGATGTGCTGCGAGATGCAGAGACGATGATCAGAAATGTGGATGC 2937
DB 2821 TCCTTTTAGATGTGCTGCGAGATGCAGAGACGATGATCAGAAATGTGGATGC 2880
QY 2938 CGTGTTCGGAATGTCCTCACACTACCAAAACAGGCGACCTTTACGACAGAGATGGAT 2997
DB 2881 CGTGTTCGGAATGTCCTCACACTACCAAAACAGGCGACCTTTACGACAGAGATGGAT 2940
QY 2998 TTGGGGTACTCTCTCCGAGGCTCAGGTGCAAGATTCTGAGAGAACAAATTTAGTTTAA 3057
DB 2941 TTGGGGTACTCTCTCCGAGGCTCAGGTGCAAGATTCTGAGAGAACAAATTTAGTTTAA 3000
QY 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTT 3117
DB 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTT 3060
QY 3118 CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAACCG 3177
DB 3061 CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGAGCG 3120
```

RESULT 12

AA03096
ID AA03096 standard; DNA; 3120 BP.

XX AC AA03096;

XX DT 14-FEB-1996 (first entry)

XX DE Protein tyrosine-kinase LprK25 gene.

XX KW Protein tyrosine-kinase; pTK; LprK25; agonist; cell growth;

XX KW differentiation; ss.

XX OS Homo sapiens.

XX PN W09527061-A1.

XX PD 12-OCT-1995.

XX PF 04-APR-1995; 95WO-US04228.

XX PR 04-APR-1994; 94US-0222616.

XX PA (GETH) GENENTECH INC.

XX PI Bennett BD, Goeddel D, Lee JW, Matthews W, Tsai SP;

XX PI Wood WJ;

XX DR WPI; 1995-366160/47.

XX AGonist antibodies which activate specific protein tyrosine

PT kinase(s) - also activate chimeric proteins of kinase extracellular

PT domain and Ig constant domain, useful for studying, and therapeutic

PT modulation of, cell growth and differentiation

XX PS Disclosure; Page 59-62; 125pp; English.

XX CC DNA probes based on protein tyrosine-kinase (pTK) sequences were used

CC to screen cDNA libraries to identify novel pTK genes. A

CC full-length sequence (AA03096) for the novel pTK LprK25 gene was

CC isolated from lymphocytic and megakaryocytic cell line libraries.

CC The gene can be used to produce recombinant LprK25, to identify other

CC new pTK genes, or to design drugs, peptides or antisense constructs

CC that modulate pTK activity.

XX SQ Sequence 3120 BP; 944 A; 642 C; 700 G; 834 T; 0 other;

Query Match 88.1%; Score 3083.2; DB 16; Length 3120;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
QY 58 ATGCGCGCGTGGCGCGCGACGGGCGACCGTGCCTGCTGCTGTTTCTTGCATG 117
DB 1 ATGAGAGCGTTGGCGCGCGACGGGCGCGCGTGCCTGCTGCTGTTTCTTGCATG 60
QY 118 ATATTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT 177
DB 61 ATATTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT 120
QY 178 AAGAACAAATGATTCATCAGTGGGAAGTCAATCATCATATCCCATGGTATCAGAAATCCCG 237
DB 121 AAGAACAAATGATTCATCAGTGGGAAGTCAATCATCATATCCCATGGTATCAGAAATCCCG 180
QY 238 GAAGACCTCGGGTGTGGCTTTGAGACCCCGAGAGCTCAGGAGACAGTGTACGAGCTCCGCT 297
DB 181 GAAGACCTCGGGTGTGGCTTTGAGACCCCGAGAGCTCAGGAGACAGTGTACGAGCTCCGCT 240
QY 298 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTGCATGCCCGAGGAGAC 357
DB 241 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTGCATGCCCGAGGAGAC 300
QY 358 ATTTCTCTCTCTGCTGCTTTAAGCACAGCTCCCTGAAATTCGACCCACATTTTGATTATA 417
DB 301 ATTTCTCTCTCTGCTGCTTTAAGCACAGCTCCCTGAAATTCGACCCACATTTTGATTATA 360
QY 418 CAAACACAGAGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGAA 477
DB 361 CAAACACAGAGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGAA 420
QY 478 TACCTACTTTTATTCAGAGTGAAGTACCAATACCAATATTCTTTACAGTGGATATA 480
DB 421 TACCTACTTTTATTCAGAGTGAAGTACCAATACCAATATTCTTTACAGTGGATATA 480
QY 538 AGAATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAC 597
DB 481 AGAATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAC 540
QY 598 GCCCTGGTCTCATATCTGAGAGCGTTCAGAGCGATCGTGGATGGTGGTTCGAT 657
DB 541 GCCCTGGTCTGCATATCTGAGAGCGTTCAGAGCGATCGTGGATGGTGGTTCGAT 600
QY 658 TCACAGGGGGAAGCTGTAAGAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTG 717
DB 601 TCACAGGGGGAAGCTGTAAGAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTG 660
QY 718 CTTCAATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAATGGSCAGGAA 777
DB 661 CTTCAATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAATGGSCAGGAA 720
QY 778 TGCACACAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTA 837
DB 721 TGCACACAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTA 780
QY 838 TTTCTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCTATGARACCATGGA 897
DB 781 TTTCTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCTATGARACCATGGA 840
QY 898 TTCGGCTCACCTGGGAATTTAGAAACAAGCACTCGAGGAGGCAACTACTTTGAGATG 957
DB 841 TTCGGCTCACCTGGGAATTTAGAAACAAGCACTCGAGGAGGCAACTACTTTGAGATG 900
QY 958 AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCATG 1017
DB 901 AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCATG 960
QY 1018 GCAAGAAACACACCGGATACACTTGTTCCTCTTCAAAGCATCCAGTCAATCAGCT 1077
DB 961 GCAAGAAACACACCGGATACACTTGTTCCTCTTCAAAGCATCCAGTCAATCAGCT 1020
QY 1078 TTGTTTACCATCGTAGGAAAGGATTATAAATGCTACCAATTTCAAGTGAAGATTATGAA 1137
DB 1021 TTGTTTACCATCGTAGGAAAGGATTATAAATGCTACCAATTTCAAGTGAAGATTATGAA 1080
QY 1138 ATTGACCAATATGAAGAGTTTCTTTTCTGTCAGGTTTAAAGCTTACCACAAATCAGA 1197
```



```
Db 1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTGAGGTTTAAAGCCTACCCACAAATCAGA 1140
Qy 1198 TGTACGTGGACCTTCTCTCGAATAATCATTTCTTGTGAGCAAAAGGGTCTTGATAACGGA 1257
Db 1141 TGTACGTGGACCTTCTCTCGAATAATCATTTCTTGTGAGCAAAAGGGTCTTGATAACGGA 1200
Qy 1258 TACAGCATATCCAAAGTTTTCGAATCATAAAGCAGCAGGAGAGATATATATTCATGCA 1317
Db 1201 TACAGCATATCCAAAGTTTTCGAATCATAAAGCAGCAGGAGAGATATATATTCATGCA 1260
Qy 1318 GAAATGATGATGCCCAATTTACCAAAATGTTACGCTGGAATATAAAGAGGAAACCTCAA 1377
Db 1261 GAAATGATGATGCCCAATTTACCAAAATGTTACGCTGGAATATAAAGAGGAAACCTCAA 1320
Qy 1378 GTGCTCGCAGAAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTACA 1437
Db 1321 GTCTCGCAGAAAGCTTCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTACA 1380
Qy 1438 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1497
Db 1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1440
Qy 1498 GGAGTCTGGAAATAGAAAGGCTTAACAGAAAGTGTGTGGACAGTGGGTCTCGAGCAGTACT 1557
Db 1441 GGAGTCTGGAAATAGAAAGGCTTAACAGAAAGTGTGTGGACAGTGGGTCTCGAGCAGTACT 1500
Qy 1558 CTAACATGAGTGAAGCCATAAAGGGTTCCTGCTCAAGTCTGTCAGTACAAATTCCTT 1617
Db 1501 CTAACATGAGTGAAGCCATAAAGGGTTCCTGCTCAAGTCTGTCAGTACAAATTCCTT 1560
Qy 1618 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1677
Db 1561 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1620
Qy 1678 AACATCTCATGTATGCAACAATTTGGTGTGTGCTCCTCTTCATTTGCTGTTTAAACCCGTG 1737
Db 1621 AACATCTCATGTATGCAACAATTTGGTGTGTGCTCCTCTTCATTTGCTGTTTAAACCCGTG 1680
Qy 1738 CTAATTTGTACAGATACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAG 1797
Db 1681 CTAATTTGTACAGATACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAG 1740
Qy 1798 GTGACCGGCTCCTCAGATAATGAGTACTTCTACCTGTGATTTACAGAGAATATGAATATGAT 1857
Db 1741 GTGACCGGATCCTCAGATTTAGTACTTCTACGTTGATTTACAGAGAATATGAATATGAT 1800
Qy 1858 CTCAAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGSTACTAGGATCAGGTGCT 1917
Db 1801 GTCAAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGSTACTAGGATCAGGTGCT 1860
Qy 1918 TTTGGAAAAGTGTAGAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAG 1977
Db 1861 TTTGGAAAAGTGTAGAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAG 1920
Qy 1978 GTTGGCGCTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
Db 1921 GTTACCGTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
Qy 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGAAATATTTGAACCTGCTGGGGGCG 2097
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGAAATATTTGAACCTGCTGGGGGCG 2040
Qy 2098 TGCACACTGTCAGGACCAATTTACTTGATTTTGAATPACTGTTGCTATGGTGATCTTCTC 2157
Db 2041 TGCACACTGTCAGGACCAATTTACTTGATTTTGAATPACTGTTGCTATGGTGATCTTCTC 2100
Qy 2158 AACTATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2217
Db 2101 AACTATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2160
Qy 2218 CACAATTTCAAGTTTAAACCCACTTTCCAAATCAGATCCAAATTCACGATGCCCTGGTTCA 2277
```

```
Db 2161 CACAATTTCAAGTTTAAACCCACTTTTCCAATCAGATCCAAATTCAGCATGCTGCTGTCA 2220
Qy 2278 AGAGAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGAAATTCATTT 2337
Db 2221 AGAGAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGAAATTCATTT 2280
Qy 2338 CACTCTGAAGATGAATTAATGAAATCAAAACCAAAAGGCTGGAAGAAGAGGAGACTTG 2397
Db 2281 CACTCTGAAGATGAATTAATGAAATCAAAACCAAAAGGCTGGAAGAAGAGGAGACTTG 2340
Qy 2398 AATGCTTTACATTTGAAGATCTTCTTTGCTTTTGCATATCAAGTTGCCAAAGAAATGGAA 2457
Db 2341 AATGCTTTACATTTGAAGATCTTCTTTGCTTTTGCATATCAAGTTGCCAAAGAAATGGAA 2400
Qy 2458 TTTCTGGAAATTAAGTCGTTTTCACAGAGACTTGGCCGCGCAGGAACTGCTTGTTCACC 2517
Db 2401 TTTCTGGAAATTAAGTCGTTTTCACAGAGACTTGGCCGCGCAGGAACTGCTTGTTCACC 2460
Qy 2518 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTTCGAGATATCATGTGATTTCC 2577
Db 2461 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTTCGAGATATCATGTGATTTCC 2520
Qy 2578 AACTATGTTGTCAGGGGCAATGCCGCTGCTGTAATAATGGATGGCCCCGCAAGACCTG 2637
Db 2521 AACTATGTTGTCAGGGGCAATGCCGCTGCTGTAATAATGGATGGCCCCGCAAGACCTG 2580
Qy 2638 TTTGAAGCATCTACACCATTAAGAGTATGATGCTGGTCTATATGGATATATCTGTTGGGAA 2697
Db 2581 TTTGAAGCATCTACACCATTAAGAGTATGATGCTGGTCTATATGGATATATCTGTTGGGAA 2640
Qy 2698 ATCTTCTCAGTTGGTGTGAATCCTTACCTGTCATTCGGTTGATGCTCAACTTCTACAAA 2757
Db 2641 ATCTTCTCAGTTGGTGTGAATCCTTACCTGTCATTCGGTTGATGCTCAACTTCTACAAA 2700
Qy 2758 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTATGTCTACAGAGAATATACATTT 2817
Db 2701 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTATGTCTACAGAGAATATACATTT 2760
Qy 2818 ATATGCAATCCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAATTTGACT 2877
Db 2761 ATAATGCAATCCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAATTTGACT 2820
Qy 2878 TCGTTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGATATCAGAAATCTGGATGCG 2937
Db 2821 TCGTTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGATATCAGAAATCTGGATGCG 2880
Qy 2938 CGTGTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2997
Db 2881 CGTGTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2940
Qy 2998 TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3057
Db 2941 TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3000
Qy 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTT 3117
Db 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTT 3060
Qy 3118 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACCAGACTTTTCTCTAGAGCGCG 3177
Db 3061 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACCAGACTTTTCTCTAGAGCGCG 3120
```

RESULT 13

AAV39041

ID AAV39041 standard; cDNA to mRNA; 2949 BP.

XX AAV39041;

XX 12-Oct-1998 (first entry)

DE Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO:24.

XX


```

1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1440
1498 GGAGCTCTGGAAATAGAAAGGCTAACAGAAAAGTGTGGACAGTGGGTGTCGAGCAGTACT 1557
1441 GGAGCTCTGGAAATAGAAAGGCTAACAGAAAAGTGTGGACAGTGGGTGTCGAGCAGTACT 1500
1558 CTAACATGAGTGAAGCCATAAAGAGGTTCCTGCTCAAGTCTGTGTCATACAATTCCTT 1617
1501 CTAACATGAGTGAAGCCATAAAGAGGTTCCTGCTCAAGTCTGTGTCATACAATTCCTT 1560
1618 GGCACATCTTGTGAGAGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1677
1561 GGCACATCTTGTGAGAGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1620
1678 AACATCTCATCTATGATCAACAATTTGGTGTGTGCTCTCTCTCTCATCTGTCGTTTAAACCTG 1737
1621 AACATCTCATCTATGATCAACAATTTGGTGTGTGCTCTCTCTCTCATCTGTCGTTTAAACCTG 1680
1738 CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAG 1797
1681 CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAG 1740
1798 GTGACCGGCTCCTCAGATAATAGTACTTCTACGTTGATTTTACAGAGATATGAATATGAT 1857
1741 GTGACCGGCTCCTCAGATAATAGTACTTCTACGTTGATTTTACAGAGATATGAATATGAT 1800
1858 CT-----CAATGGGAGTTTCCAAGAGAAAATTTA 1887
1801 CTCAAAATGGGAGTTTCCAAGAGAAAATTTGGCACAAATGGGAGTTTCCAAGAGAAAATTTA 1860
1888 GAGTTTGGGAGGTACTAGATCAGGTGCTTTTGGAAAAGTGAAGCAACGACAGCTTAT 1947
1861 GAGTTTGGGAGGTACTAGATCAGGTGCTTTTGGAAAAGTGAAGCAACGACAGCTTAT 1920
1948 GGAATTAGCAAAAACAGAGAGTCTCAATCCAGGTTCCCGTCAAAATGCTGAAAGAAAAGCA 2007
1921 GGAATTAGCAAAAACAGAGAGTCTCAATCCAGGTTCCCGTCAAAATGCTGAAAGAAAAGCA 1980
2008 GACAGCTCTGAAGAGAGGAGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGC 2067
1981 GACAGCTCTGAAGAGAGGAGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGC 2040
2068 CACGAGAAATTTGTGACCTGCTGGGGCGTGCACACTGTCAGACCACTTACTTCTGATT 2127
2041 CACGAGAAATTTGTGACCTGCTGGGGCGTGCACACTGTCAGACCACTTACTTCTGATT 2100
2128 TTTGAATCTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAATTT 2187
2101 TTTGAATCTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAATTT 2160
2188 CACAGGACTTGGACAGAGATTTTCAAGAGACACAAATTTACGTTTTTACCCGACCTTCCAA 2247
2161 CACAGGACTTGGACAGAGATTTTCAAGAGACACAAATTTACGTTTTTACCCGACCTTCCAA 2220
2248 TCACATCCAAATTCACAGCTGCTGTTCAAGAGAGTTCAGATACACCCGAGCTCGGAT 2307
2221 TCACATCCAAATTCACAGCTGCTGTTCAAGAGAGTTCAGATACACCCGAGCTCGGAT 2280
2308 CAAATCTCAGGGCTTCATGGGAATTTCACTCTGAAAGTGAATTTGAATATGAAAAC 2367
2281 CAAATCTCAGGGCTTCATGGGAATTTCACTCTGAAAGTGAATTTGAATATGAAAAC 2340
2368 CAAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTTCTTTGC 2427
2341 CAAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTTCTTTGC 2400
2428 TTTGCATATCAAGTTGCCAAGGAATTCGAATTTCTGGAATTTAAGTGTGTTGTTACAG 2487
2401 TTTGCATATCAAGTTGCCAAGGAATTTCTGGAATTTAAGTGTGTTGTTGTTACAG 2460
2488 GACCTGGCCGACGAGGAGCTGTTGTCACCCAGGAAAGTGGTGAAGATATGACATTT 2547
```

```

2461 GACCTGGCCGCCAGGAACGTGCTTGTCAACCACGGAAGAGTGTGAAGATATGTGACTTT 2520
2548 GGATTGGCTCGAGATATCATGAGTATATCCAACTATGTTGTGTCAGGGGCAATGCCGCTGTG 2607
2521 GGATTGGCTCGAGATATCATGAGTATATCCAACTATGTTGTGTCAGGGGCAATGCCGCTGTG 2580
2608 CCTGTAAATGAGTGGCCCGGAAAGCCGTGTTTGAAGGCATCTACACCATTAAGAGTGAT 2667
2591 CCTGTAAATGAGTGGCCCGGAAAGCCGTGTTTGAAGGCATCTACACCATTAAGAGTGAT 2640
2668 GTCTGGTCATATGGAATATTAATCTGTTGGAAAATCTTCTCACTTGGTGTGTAATCCTTACCCT 2727
2641 GTCTGGTCATATGGAATATTAATCTGTTGGAAAATCTTCTCACTTGGTGTGTAATCCTTACCCT 2700
2728 GGCATTCGGTGTGATGCTAACTCTACAAAAGTGAATCAAAATGGATTTAAAATGGATCAG 2787
2701 GGCATTCGGTGTGATGCTAACTCTACAAAAGTGAATCAAAATGGATTTAAAATGGATCAG 2760
2788 CCATTTTATGTACAGAAAATATACATTAATGCAATCTCTGCTGGCTTTTGAATCTCA 2847
2761 CCATTTTATGTACAGAAAATATACATTAATGCAATCTCTGCTGGCTTTTGAATCTCA 2820
2848 AGGAAACGGCCATCCTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTGGCAGATGCA 2907
2821 AGGAAACGGCCATCCTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTGGCAGATGCA 2880
2908 GAAGAAGCGATGTATCAGAAATGTGATGCGCGTGTTCGGAATGTCCCTCACACCTACCAA 2967
2881 GAAGAAGCGATGTATCAGAAATGTGATGCGCGTGTTCGGAATGTCCCTCACACCTACCAA 2940
2968 AACAGGCGA 2976
2941 AACAGGCGA 2949

RESULT 14
AAV39040
ID AAV39040 standard; cDNA to mRNA; 2958 BP.
XX AC AAV39040;
XX XX
DT 12-OCT-1998 (first entry)
XX
DE Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO:23.
DE XX
KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
KW membrane-vicinal region; tyrosine kinase; juxtamembrane region; ds.
XX OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2958
FT FT /*tag= a
FT FT /note= "no stop codon given"
XX
PN W09817808-A1.
XX
PD 30-APR-1998.
XX
PF 13-OCT-1997; 97WO-JP03667.
XX
PR 18-OCT-1996; 96JP-0297329.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
PI Yokota S;
XX
DR WPI; 1998-362333/31.
XX P-PSDB; AAW63587.
XX
PT Nucleic acid sequences encoding receptor type protein kinase -
M2 useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
```


Db	1801	CTCAAAATGGGAGTTTGGATTTTCAGAGAAATAGAAATATGATATCAAAATGGGAGTTTCCAAAGA	1860
Qy	1879	GAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTTGGAAAAGATGATGAAGCGCA	1938
Db	1861	GAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGTAGCAAGCGCA	1920
Qy	1939	ACAGCTTATGAAATTAGCAAAACAGGAGTCTCAATCCAGGTGGCGTCAAAAATGCTGAAA	1998
Db	1921	ACAGCTTATGAAATTAGCAAAACAGGAGTCTCAATCCAGGTGGCGTCAAAAATGCTGAAA	1980
Qy	1999	GAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAACTCAAGATGATGACCCAG	2058
Db	1981	GAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAACTCAAGATGATGACCCAG	2040
Qy	2059	CTGGGAAGCCACGAGAAATATTTGTGAACCTGCTGGGGCGTGCACTGTTCAGGACCAATTT	2118
Db	2041	CTGGGAAGCCACGAGAAATATTTGTGAACCTGCTGGGGCGTGCACTGTTCAGGACCAATTT	2100
Qy	2119	TACTTTGAATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAGTAAAGA	2178
Db	2101	TACTTTGAATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAGTAAAGA	2160
Qy	2179	GAAAAATTTTCACAGGACTTGCACAGAGATTTTCAAGGAAACACAATTTTCAGTTTATACCC	2238
Db	2161	GAAAAATTTTCACAGGACTTGCACAGAGATTTTCAAGGAAACACAATTTTCAGTTTATACCC	2220
Qy	2239	ACTTTCCAATCACATCCAAATTCACGATGCCTGGTTCAAGAGAAGTTTCAGATACACCCG	2298
Db	2221	ACTTTCCAATCACATCCAAATTCACGATGCCTGGTTCAAGAGAAGTTTCAGATACACCCG	2280
Qy	2299	GACTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTCACCTCTCAAGATGAAATGAA	2358
Db	2281	GACTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTCACCTCTCAAGATGAAATGAA	2340
Qy	2359	TATGAAACCAAAAAAGGCTTGAAGAAGAGAGGACTTGAATCTGCTTACATTTGAAGAT	2418
Db	2341	TATGAAACCAAAAAAGGCTTGAAGAAGAGAGGACTTGAATCTGCTTACATTTGAAGAT	2400
Qy	2419	CTTCTTTGCTTTGCATATCAATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT	2478
Db	2401	CTTCTTTGCTTTGCATATCAATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT	2460
Qy	2479	GTTCACAGAGACCTGGCCGCCAGAGAGCTGTTGCACCCACGGGAAAGTGTGGAAGATA	2538
Db	2461	GTTCACAGAGACCTGGCCGCCAGAGAGCTGTTGCACCCACGGGAAAGTGTGGAAGATA	2520
Qy	2539	TGTGACTTTGGATTGGCTCGAGATATCATCAGATGATTCACAATATCTGTCAGGGCAAT	2598
Db	2521	TGTGACTTTGGATTGGCTCGAGATATCATCAGATGATTCACAATATCTGTCAGGGCAAT	2580
Qy	2599	GCOCGTCTGCTGTAAATGGATGGCCCCCGAAAGCCTGTTTGAAGGCATCTACACCAAT	2658
Db	2581	GCOCGTCTGCTGTAAATGGATGGCCCCCGAAAGCCTGTTTGAAGGCATCTACACCAAT	2640
Qy	2659	AAGAGTATGTCGTGTCATATGGAATATTTACTGTGGAAATCTTCTCACTTGGTGTGAAT	2718
Db	2641	AAGAGTATGTCGTGTCATATGGAATATTTACTGTGGAAATCTTCTCACTTGGTGTGAAT	2700
Qy	2719	CTTTACCTCGGCATTCGGTTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAA	2778
Db	2701	CTTTACCTCGGCATTCGGTTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAA	2760
Qy	2779	ATGGATCAGCCATTTTATGCTACAGAGAATAATACATTTATATGCAATCCCTGCGGCT	2838
Db	2761	ATGGATCAGCCATTTTATGCTACAGAGAATAATACATTTATATGCAATCCCTGCGGCT	2820
Qy	2839	TTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTG	2898
Db	2821	TTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTG	2880
Qy	2899	GCAGATGCAGAGAAGCGGATGATCAGATGTTGGATGGCCGTGTTTCGGAAATGCTCTAC	2958

Db	2851	GCAGATGCAGAAAGACGATGTATCAGAAATGTGGATGCCCGTGTTCGGAATGTCTCTCAC	2940
Qy	2959	ACCTACCAAAACAGCGCA	2976
Db	2941	ACCTACCAAAACAGCGCA	2958
RESULT 15			
ID	AAV39042	standard; cDNA to mRNA; 2958 BP.	
XX	AAV39042;		
XX	12-OCT-1998	(first entry)	
XX	Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO:25.		
XX	Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;		
KW	membrane-vincinal region; tyrosine kinase; juxtamembrane region; ds.		
XX	Homo sapiens.		
OS			
XX	Key	Location/Qualifiers	
EH	CDS	1..2958	
FT		/*tag= a	
FT		/note= "no stop codon given"	
XX			
PN	W09817808-A1.		
XX	30-APR-1998.		
PD			
XX	13-OCT-1997;	97WO-JP03667.	
PF			
XX	18-OCT-1996;	96JP-0297329.	
PR			
XX	(TAKI) TAKARA SHUZO CO LTD.		
PA			
XX	Yokota S;		
PI			
XX	WPI: 1998-362333/31.		
DR	P-PSDB; AAW63589.		
XX			
PT	Nucleic acid sequences encoding receptor type protein kinase -		
PT	useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes		
PT	M2		
XX			
PS	Claim 7; Page 65-66; 80pp; Japanese.		
XX			
CC	New nucleic acid sequences have been isolated which encode receptor type		
CC	protein kinases (especially a tyrosine kinases) having tandem repeats in		
CC	the juxtamembrane region. Also described in the present invention are:		
CC	(1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)		
CC	representing FLT3 juxtamembrane receptor type protein kinases found in		
CC	leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences		
CC	encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies		
CC	recognising the kinases or their portions including the tandem repeat		
CC	region; (3) nucleic acid sequences hybridising with the nucleic acids;		
CC	(4) a method for detecting the nucleic acid sequences in human tissue		
CC	samples by: (i) extracting nucleic acid; (ii) amplifying the kinase		
CC	gene, and (iii) determining the size of the gene for comparing the size		
CC	of the normal gene not containing tandem repeats, and (5) kits for		
CC	carrying out the detection. The products and methods may be used for		
CC	diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.		
XX			
SQ	Sequence 2958 BP; 893 A; 606 C; 671 G; 788 T; 0 other;		
Query Match 81.9%; Score 2868.4; DB 19; Length 2958;			
Best Local Similarity 98.6%; Pred. No. 0;			
Matches 2918; Conservative 0; Mismatches 1; Indels 39; Gaps 1;			
Qy	58	ATCCCGCGGTGCGCGGACGCGGACCGTGCCTGCTGTTGTTTCTGCAAG 117	
Db	1	ATCCCGCGGTGCGCGGACGCGGACCGTGCCTGCTGTTGTTTCTGCAAG 60	

Db	2221	ACTTTCCAAATCACATCCAAATTCAGCATGCCGTTCAGAGAAAGTTTCAGATACACCG	2280
Qy	2299	GACTCGGATCAAACTCTAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATTGAA	2358
Db	2281	GACTCGGATCAAACTCTAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATTGAA	2340
Qy	2359	TATGAAACCAAAAGGCTGGAAGAGAGAGGACTTCAATGTGCTTACATTTGAAGAT	2418
Db	2341	TATGAAACCAAAAGGCTGGAAGAGAGAGGACTTGAATGTGCTTACATTTGAAGAT	2400
Qy	2419	CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTG	2478
Db	2401	CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTG	2460
Qy	2479	GTTCCACAGACCTGGCCGCCAGGAACGCTGTGTCACCCAGGGAAGTGTGAAGATA	2538
Db	2461	GTTCCACAGACCTGGCCGCCAGGAACGCTGTGTCACCCAGGGAAGTGTGAAGATA	2520
Qy	2539	TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTGTACAGGGCAAT	2598
Db	2521	TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTGTACAGGGCAAT	2580
Qy	2599	GCCCGTCTGCTGTAATAATGGATGCCCCCGAAAGCCCTGTTTGAAGGCATCTACACCAT	2658
Db	2581	GCCCGTCTGCTGTAATAATGGATGCCCCCGAAAGCCCTGTTTGAAGGCATCTACACCAT	2640
Qy	2659	AAGAGTGATGCTGGTCATATGGAATATTAATCTGTTGGGAAATCTTCTCACTTGGTGTGAAT	2718
Db	2641	AAGAGTGATGCTGGTCATATGGAATATTAATCTGTTGGGAAATCTTCTCACTTGGTGTGAAT	2700
Qy	2719	CTTTACCTCGCATTCGGTGTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAAA	2778
Db	2701	CTTTACCTCGCATTCGGTGTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAAA	2760
Qy	2779	ATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATAATGCAATCCTGCTGGCT	2838
Db	2761	ATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATAATGCAATCCTGCTGGCT	2820
Qy	2839	TTTGACTCAAGGAACGGCCATCCTTCCTTAATTTGACTTCGTTTTTATAGGATGTCAGCTG	2898
Db	2821	TTTGACTCAAGGAACGGCCATCCTTCCTTAATTTGACTTCGTTTTTATAGGATGTCAGCTG	2880
Qy	2899	GCAGATGCAGAGAACGCGATGATCAGAAATGTTGGATGCCGCTGTTTCGGAATGTCCTCAC	2958
Db	2881	GCAGATGCAGAGAACGCGATGATCAGAAATGTTGGATGCCGCTGTTTCGGAATGTCCTCAC	2940
Qy	2959	ACCTACCAAAACAGGCGA	2976
Db	2941	ACCTACCAAAACAGGCGA	2958

Search completed: May 25, 2003, 12:15:03
Job time : 710.832 secs


```

, APPLICATION NUMBER: US 07/679,666
, FILING DATE: 02-APR-1991
, ATTORNEY/AGENT INFORMATION:
, NAME: Feit, Irving N.
, REGISTRATION NUMBER: 28,601
, REFERENCE/DOCKET NUMBER: LEM-3-7P
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 212-645-1405
, TELEFAX: 212-645-2054
, INFORMATION FOR SEQ ID NO: 3:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 3501 base pairs
, TYPE: NUCLEIC ACID
, STRANDEDNESS: double
, TOPOLOGY: linear
, MOLECULE TYPE: cDNA
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
, FRAGMENT TYPE: N-terminal
, FEATURE:
, NAME/KEY: CDS
, LOCATION: 58..3039
, FEATURE:
, NAME/KEY: mat_peptide
, LOCATION: 139..3036
, FEATURE:
, NAME/KEY: sig_peptide
, LOCATION: 58..138
, US-07-977-451-3

```

Query Match	100.0%	Score 3501;	DB 1;	Length 3501;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 3501:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGAGCGGCATCCAGAGCTGGCGCGCCCTGGGGACCCCGGCTCCGAGGCCATG	60	
DB	1	CGAGCGGCATCCAGAGCTGGCGCGCCCTGGGGACCCCGGCTCCGAGGCCATG	60	
QY	61	CCGCGGTTGGCGCGCGAGCGCGGGACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA	120	
DB	61	CCGCGGTTGGCGCGCGAGCGCGGGACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA	120	
QY	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAG	180	
DB	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAG	180	
QY	181	AACAATGATTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGGAA	240	
DB	181	AACAATGATTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGGAA	240	
QY	241	GACCTCGGGTGTGGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300	
DB	241	GACCTCGGGTGTGGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300	
QY	301	GAAGTGATGTATCTGCTTCATCATCACTGCAAGTGTGTCGATGCCCGAGGAACATT	360	
DB	301	GAAGTGATGTATCTGCTTCATCATCACTGCAAGTGTGTCGATGCCCGAGGAACATT	360	
QY	361	TCCTGCTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420	
DB	361	TCCTGCTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420	
QY	421	AACAGAGGAGTTGTTTCATGGTCATTTTGAATAATGACAGAAACCCAGCTGGAGAAATC	480	
DB	421	AACAGAGGAGTTGTTTCATGGTCATTTTGAATAATGACAGAAACCCAGCTGGAGAAATC	480	
QY	481	CTACTTTTTATTTCAGAGTGAAGCTACCAATTTACAAATATTGTTTACAGTGAATAAGA	540	
DB	481	CTACTTTTTATTTCAGAGTGAAGCTACCAATTTACAAATATTGTTTACAGTGAATAAGA	540	
QY	541	AATACCCCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAAATGGAAACACAGGACGCC	600	
DB	541	AATACCCCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAAATGGAAACACAGGACGCC	600	

QY	601	CTGCTCTGCATATCTCAGAGCGTTCAGAGCCGATCGTGGATGGTGCTTTGGCAATTC	660
Db	601		
QY	661	CTGCTCTGCATATCTCAGAGCGTTCAGAGCCGATCGTGGATGGTGCTTTGGCAATTC	720
Db	661		
QY	721	CAGGGGAAAGCTGTTAAAGAAAGTCCAGTGTGTTTAAAGAGGAGAAAAGTGCTT	780
Db	721		
QY	781	CAGGGGAAAGCTGTTAAAGAAAGTCCAGTGTGTTTAAAGAGGAGAAAAGTGCTT	840
Db	781		
QY	841	CATGAATATTTCGGACGACATAAAGTGCTGCCAGAAATGAACCTGGCGCAGGAATGC	900
Db	841		
QY	901	CATGAATATTTCGGACGACATAAAGTGCTGCCAGAAATGAACCTGGCGCAGGAATGC	960
Db	901		
QY	961	ACCAAGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACACCATTCGCACAATATTT	1020
Db	961		
QY	1021	ACCAAGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACACCATTCGCACAATATTT	1080
Db	1021		
QY	1081	AGAAACGACACCCGGATACACTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG	1140
Db	1081		
QY	1141	GTTACCATCGTAGGAAGGGATTATAAATGCTACCAATTCAGTGAAGATTATGAAAT	1200
Db	1141		
QY	1201	GTTACCATCGTAGGAAGGGATTATAAATGCTACCAATTCAGTGAAGATTATGAAAT	1260
Db	1201		
QY	1261	ACGTGGACCTTCTCGAAAATCATTTCCCTGTGAGCAAAAGGCTTTGATACGGATAC	1320
Db	1261		
QY	1321	ACGTGGACCTTCTCGAAAATCATTTCCCTGTGAGCAAAAGGCTTTGATACGGATAC	1380
Db	1321		
QY	1381	AGCATATCCAAAGTTTTCGAATCAATAGCAACCCAGGAGAAATATATTTCCATGCGAA	1440
Db	1381		
QY	1441	AATGATGATGCCAATTTTACCAAAATGTTTCCAGCTGAATATAGAAGAAACCTCAAGT	1500
Db	1441		
QY	1501	AATGATGATGCCAATTTTACCAAAATGTTTCCAGCTGAATATAGAAGAAACCTCAAGT	1560
Db	1501		
QY	1561	CTCGCAGAAGCATCGGCAAGTCAGGCGTCCTGTTTCTCGGATGGATACCCATTTACCAT	1620
Db	1561		
QY	1621	CTCGCAGAAGCATCGGCAAGTCAGGCGTCCTGTTTCTCGGATGGATACCCATTTACCAT	1680
Db	1621		
QY	1681	TGGACCTGGAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAAGATTCACAGAAG	1740
Db	1681		
QY	1741	TGGACCTGGAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAAGATTCACAGAAG	1800
Db	1741		
QY	1801	GTCTGGAATAGAAAGGCTTAAAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA	1860
Db	1801		
QY	1861	GTCTGGAATAGAAAGGCTTAAAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA	1920
Db	1861		
QY	1921	RACATGAGTGAAGCCATAAAAGGTTTCCCTGTGATCAAAATTCCTTTGGC	1980
Db	1921		
QY	1981	RACATGAGTGAAGCCATAAAAGGTTTCCCTGTGATCAAAATTCCTTTGGC	2040
Db	1981		
QY	2041	ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTCAATCCAGACAAC	2100
Db	2041		
QY	2101	ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTCAATCCAGACAAC	2160
Db	2101		

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
US-08-252-517-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGAGGGCGCATCCGAGGGCTGGCGCGGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGGCGCATCCGAGGGCTGGCGCGGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
Qy 61 CGCGGCTTGGCGCGGACCGGCGACCCGTGCGCGTCTGTTGTTTTCGCAATGATA 120
Db 61 CGCGGCTTGGCGCGGACCGGCGACCCGTGCGCGTCTGTTGTTTTCGCAATGATA 120
Qy 121 TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Db 121 TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
Qy 181 AACATGATTTCATCAGTGGGGAAGTCATCATATCCCATGATGATCCAGAAATCCCGGAA 240
Db 181 AACATGATTTCATCAGTGGGGAAGTCATCATATCCCATGATGATCCAGAAATCCCGGAA 240
Qy 241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGAGTGTAGGAACTGCCCGTGTG 300
Db 241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGAGTGTAGGAACTGCCCGTGTG 300
Qy 301 GAAGTGGATGATCTGCTTCCATCAGTCAAGTGTGCTGATGATGATGATGATGATGATGAT 360
Db 301 GAAGTGGATGATCTGCTTCCATCAGTCAAGTGTGCTGATGATGATGATGATGATGATGAT 360
Qy 361 TCCTGTCTCTGGTCTTTAAGCAGAGCTCCCTGAAATGTCAGCCACATTTTGAATTTACAA 420
Db 361 TCCTGTCTCTGGTCTTTAAGCAGAGCTCCCTGAAATGTCAGCCACATTTTGAATTTACAA 420
Qy 421 AACAGAGGAGTGTGTTTCCATGCTCAATTTGAAATGACAGAAACCCCAAGCTGGAGAAATAC 480
Db 421 AACAGAGGAGTGTGTTTCCATGCTCAATTTGAAATGACAGAAACCCCAAGCTGGAGAAATAC 480
Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATACAAATATGTTTACAGTGAATATAAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATACAAATATGTTTACAGTGAATATAAGA 540
Qy 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACCCAGGACGCC 600
Db 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACCCAGGACGCC 600
Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTGTTCCGATTCA 660
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTGTTCCGATTCA 660
Qy 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720
Qy 721 CATGAATTTTGGGACGACATAAGTGTGTCAGAGAAATGAACCTGGCAGGGAATGC 780
Db 721 CATGAATTTTGGGACGACATAAGTGTGTCAGAGAAATGAACCTGGCAGGGAATGC 780
Qy 781 ACCAGGCTGTTCCAAATAGATCTAAATCAAACTCCCTCAGACCACTTGCCACAATATTT 840
Db 781 ACCAGGCTGTTCCAAATAGATCTAAATCAAACTCCCTCAGACCACTTGCCACAATATTT 840
Qy 841 CTTAAAGTAGGGGAACCTTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
Db 841 CTTAAAGTAGGGGAACCTTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
Qy 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCAGAGAGGGCACTACTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCAGAGAGGGCACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCGTAGGAAAGGATTTAAATGCTACCAATTTCAAGTGAAGATTTGAAT 1140
Db 1081 GTTACCATCGTAGGAAAGGATTTAAATGCTACCAATTTCAAGTGAAGATTTGAAT 1140
Qy 1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATG 1200
Db 1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATG 1200
Qy 1201 ACCTGGACCTTCTCTGAAATCATTTCCCTGTGTGAGCAAAAGGCTCTTGATAACGGATAC 1260
Db 1201 ACCTGGACCTTCTCTGAAATCATTTCCCTGTGTGAGCAAAAGGCTCTTGATAACGGATAC 1260

1261 AGCATATCCAAAGTTTGCATCATAAAGCACCAGCCAGGAGAAATATATATCCATGTCAGAA 1320
1261 AGCATATCCAAAGTTTGCATCATAAAGCACCAGCCAGGAGAAATATATATCCATGTCAGAA 1320
1321 AATGATGATGCCAAATTTACCAAAATGTTTCCAGCTGATATATAAGAGAAACCTCAAGTG 1380
1321 AATGATGATGCCAAATTTACCAAAATGTTTCCAGCTGATATATAAGAGAAACCTCAAGTG 1380
1381 CTCGAGAGAGCATCGCAAGTCAGGCGTCTGTTCTCGGATGGATACCCATACCATCT 1440
1381 CTCGAGAGAGCATCGCAAGTCAGGCGTCTGTTCTCGGATGGATACCCATACCATCT 1440
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCCTCCAACTGCACAGAGAGATCACAGAAGGA 1500
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCCTCCAACTGCACAGAGAGATCACAGAAGGA 1500
1501 GTCGGAATAGAAAGGCTAAACAGAAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
1501 GTCGGAATAGAAAGGCTAAACAGAAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620
1621 ACATCTGTGAGAGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTATCCCAAGACAAC 1680
1621 ACATCTGTGAGAGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTATCCCAAGACAAC 1680
1681 ATCTCATCTATGCAACAATTTGGTGTGCTCTCTCTTCTATTTGCTGTTTAAACCTGCTA 1740
1681 ATCTCATCTATGCAACAATTTGGTGTGCTCTCTCTTCTATTTGCTGTTTAAACCTGCTA 1740
1741 ATTTGTCAGAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCAGAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCTCAGATAATGAGTACTCTAGCTTGCATTTTCAGAGAAATATGAATATGATCTC 1860
1801 ACCGGCTCTCAGATAATGAGTACTCTAGCTTGCATTTTCAGAGAAATATGAATATGATCTC 1860
1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1921 GGAAGTGTGACGCAACAGCTTTAGTAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 GGAAGTGTGACGCAACAGCTTTAGTAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 GCCGTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040
1981 GCCGTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGGTGC 2100
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGGTGC 2100
2101 ACATGTGACGACCAATTTACTTGAATTTTGAATGATGTTGCTATGTTGATCTCTCAAC 2160
2101 ACATGTGACGACCAATTTACTTGAATTTTGAATGATGTTGCTATGTTGATCTCTCAAC 2160
2161 TATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2221 AATTTCAAGTTTATACCCACTTTCCAAATCATCCAAATTCACAGATGCCGTGTTCAAGA 2280
2221 AATTTCAAGTTTATACCCACTTTCCAAATCATCCAAATTCACAGATGCCGTGTTCAAGA 2280
2281 GAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
2281 GAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340

2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400
2401 GTGCTTACATTTGAAGATCTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2401 GTGCTTACATTTGAAGATCTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2461 CTGGAATTTAAGTCGTCTGTTTCACAGAGACTGGCCGCCAGCAAGCTGTTGTACCCAC 2520
2461 CTGGAATTTAAGTCGTCTGTTTCACAGAGACTGGCCGCCAGCAAGCTGTTGTACCCAC 2520
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCAC 2580
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCAC 2580
2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCGCCGCAAGCCTGTTT 2640
2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCGCCGCAAGCCTGTTT 2640
2641 GAAGCATCTACACCATTAAGAGTGTCTGCTGTCATATGGAATATTTACTGTGGGAATC 2700
2641 GAAGCATCTACACCATTAAGAGTGTCTGCTGTCATATGGAATATTTACTGTGGGAATC 2700
2701 TTCTCACTTGTGTGAATCTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACTG 2760
2701 TTCTCACTTGTGTGAATCTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACTG 2760
2761 ATTCAAAATGGAATTTAAATGGATGAGCATTTTATGCTACAGAGAATATACATATA 2820
2761 ATTCAAAATGGAATTTAAATGGATGAGCATTTTATGCTACAGAGAATATACATATA 2820
2821 ATGCAATCTCTCTGGGCTTTTACTCAAGAAAAGGCCATCTTCCCTAAATTTGACTTCG 2880
2821 ATGCAATCTCTCTGGGCTTTTACTCAAGAAAAGGCCATCTTCCCTAAATTTGACTTCG 2880
2881 TTTTATGATCTCAGCTGGCAGATGCAGAAAGGAGTATATCAAAATGTTGGATGGCCT 2940
2881 TTTTATGATCTCAGCTGGCAGATGCAGAAAGGAGTATATCAAAATGTTGGATGGCCT 2940
2941 GTTTCCGGAATGTCCTACACCTACAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
2941 GTTTCCGGAATGTCCTACACCTACAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
3001 GGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
3001 GGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTTCA 3120
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTTCA 3120
3121 CACTAAAAGAAAATCTATTATCACTGCTTCCACAGACTTTTCTCTAGAGCGCTCT 3180
3121 CACTAAAAGAAAATCTATTATCACTGCTTCCACAGACTTTTCTCTAGAGCGCTCT 3180
3181 CGGTTTACTCTGTTTCCAAAGGAGCTTTTGTAAAATCAAAATCATCTCTGTCACAGGCA 3240
3181 CGGTTTACTCTGTTTCCAAAGGAGCTTTTGTAAAATCAAAATCATCTCTGTCACAGGCA 3240
3241 GAGGAGCTGATTAATGAACCTTTATTCGAGCATTTGATCTGCATCCCAAGGCTTCTCAGGCG 3300
3241 GAGGAGCTGATTAATGAACCTTTATTCGAGCATTTGATCTGCATCCCAAGGCTTCTCAGGCG 3300
3301 GCTTGAAGTGAATTTGTACCTGAGTACAGTATATTTCTTGAATATACATAAAACAAAAGC 3360
3301 GCTTGAAGTGAATTTGTACCTGAGTACAGTATATTTCTTGAATATACATAAAACAAAAGC 3360
3361 ATTTTGTGAGAGAGCAATATGATTTTAAAGTCTGATGTTTAAATATATATGTA 3420
3361 ATTTTGTGAGAGAGCAATATGATTTTAAAGTCTGATGTTTAAATATATATGTA 3420
3421 TTTTCACTATTTAGTATATATTTTATGGGTGGGAATAAATTTCTACTACAGAAA 3480

||||| 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
||||| 3481 AAAAAAAAAAAAAAAAAAAAA 3501
||||| 3481 AAAAAAAAAAAAAAAAAAAAA 3501

Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
QY 3481 AAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

RESULT 3
US-07-906-397A-3
; Sequence 3, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TORIPORENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3501 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..3039
US-07-906-397A-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCGGCATCCGAGGGCTGGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
||||| 1 CGAGCGGCATCCGAGGGCTGGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGCGGCATCCGAGGGCTGGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60

QY 1141 GACCAATATGAAGAGTTTTGTTTTCTGTGTCAGGTTTAAAGCCTACCACAAATCAGATGT 1200
DB 1141 GACCAATATGAAGAGTTTTGTTTTCTGTGTCAGGTTTAAAGCCTACCACAAATCAGATGT 1200
QY 1201 ACCTGGACCTTCTCTCGAAATCATTTCCCTGTGTGAGCAAAAGGCTCTGTATACGGATAC 1260
DB 1201 ACCTGGACCTTCTCTCGAAATCATTTCCCTGTGTGAGCAAAAGGCTCTGTATACGGATAC 1260
QY 1261 AGCATATCCAAAGTTTTCATATGAACACAGCCAGGAGAAATATATATTCATTCAGAGAA 1320
DB 1261 AGCATATCCAAAGTTTTCATATGAACACAGCCAGGAGAAATATATATTCATTCAGAGAA 1320
QY 1321 AATGATGATCCCAATTTACCAAAATGTTTCCAGCTGAATATAAGAAAGAAACCTTCAAGTG 1380
DB 1321 AATGATGATCCCAATTTACCAAAATGTTTCCAGCTGAATATAAGAAAGAAACCTTCAAGTG 1380
QY 1381 CTGCGAGAAGCATCGGCAAGTCAGGCTCCTGTTTCTCGATGGATACCCATTTACCATCT 1440
DB 1381 CTGCGAGAAGCATCGGCAAGTCAGGCTCCTGTTTCTCGATGGATACCCATTTACCATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500
QY 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTCAGAGTGGTGTGAGCAGTACTCTA 1560
DB 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTCAGAGTGGTGTGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGTCAGTCAAGTCTGTCATCAATTCCTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGTCAGTCAAGTCTGTCATCAATTCCTTGGC 1620
QY 1621 ACATCTGTGAGAGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAGAAC 1680
DB 1621 ACATCTGTGAGAGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAGAAC 1680
QY 1681 ATCTCATCTATCAACAAATGTTGTGTCCTCTCTCATTTGTCGTTTAAACCTGCTA 1740
DB 1681 ATCTCATCTATCAACAAATGTTGTGTCCTCTCTCATTTGTCGTTTAAACCTGCTA 1740
QY 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
QY 1801 ACCGGCTCTCAGATATGAGTACTTCTAGTTGATTTTCAGAGATATGATATGATCTC 1860
DB 1801 ACCGGCTCTCAGATATGAGTACTTCTAGTTGATTTTCAGAGATATGATATGATCTC 1860
QY 1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
DB 1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGAGTATGACGCAACAGCTTTAGGAATAGCAAAACAGGAGTCTCAATCCAGTT 1980
DB 1921 GGAAGAGTATGACGCAACAGCTTTAGGAATAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGACCTCATGTCAGAA 2040
DB 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGACCTCATGTCAGAA 2040
QY 2041 CTCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGCGTGC 2100
DB 2041 CTCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGCGTGC 2100
QY 2101 ACAGTGTAGGACCAATTTACTGATTTTGAATGCTATGCTATGCTATGCTTCTCAAC 2160
DB 2101 ACAGTGTAGGACCAATTTACTGATTTTGAATGCTATGCTATGCTATGCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTTCAGTTTTTACCCCACTTTCCCAATCATCAATCCAAATTCACGATGCCTGGTTCAGAA 2280

DB 2221 AATTTTCAGTTTTTACCCCACTTTCCAAATCATCAATCCAAATTCAGCATGCCTGGTTCAGAA 2280
QY 2281 GAACTTCAGATACACCCGGAGTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340
DB 2281 GAACTTCAGATACACCCGGAGTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGCGTGGGAAGAGAGGAGACTTGAAT 2400
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGCGTGGGAAGAGAGGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTATCAAGTTGCCAAAGGAATGGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTATCAAGTTGCCAAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCACAGAACCTGGCCGCCAGGAAGTGTCTTCACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCACAGAACCTGGCCGCCAGGAAGTGTCTTCACCCAC 2520
QY 2521 GGGAAAGTGTGAAGATATGTGACTTTGGCTCGAGATATCATGATGATTCACCAAC 2580
DB 2521 GGGAAAGTGTGAAGATATGTGACTTTGGCTCGAGATATCATGATGATTCACCAAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCCCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640
DB 2581 TATGTTGTGAGGGCAATGCCCCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640
QY 2641 GAAGCATCTACACCATTAAGAGTATGATGCTGCTATGGAATATTTACTGTGGGAAATC 2700
DB 2641 GAAGCATCTACACCATTAAGAGTATGATGCTGCTATGGAATATTTACTGTGGGAAATC 2700
QY 2701 TTCTCAGTTGTGTGAATCCTTACCTGGCATTCGGTGTGATGCTAACTTCTACAACTG 2760
DB 2701 TTCTCAGTTGTGTGAATCCTTACCTGGCATTCGGTGTGATGCTAACTTCTACAACTG 2760
QY 2761 ATTCAAAATGATTTAAAATGATCAGCCATTTTATGCTACAGAGAATAATACATTATA 2820
DB 2761 ATTCAAAATGATTTAAAATGATCAGCCATTTTATGCTACAGAGAATAATACATTATA 2820
QY 2821 ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTCCCTAAATTTGACTCG 2880
DB 2821 ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTCCCTAAATTTGACTCG 2880
QY 2881 TTTTATGATGTCAGCTGGCAGATGCAGAAAGGATGATACAAATGTCGATGCCGT 2940
DB 2881 TTTTATGATGTCAGCTGGCAGATGCAGAAAGGATGATACAAATGTCGATGCCGT 2940
QY 2941 GTTTCGGAAATCTCTCACACCTTACCAACAGGCGACCTTTCAGCAGAGAGATGGATTG 3000
DB 2941 GTTTCGGAAATCTCTCACACCTTACCAACAGGCGACCTTTCAGCAGAGAGATGGATTG 3000
QY 3001 GGGTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTTAAG 3060
DB 3001 GGGTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTTAAG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTCA 3120
DB 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTCA 3120
QY 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTACCAGACTTTTCTCTAGAGCGCTCT 3180
DB 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTACCAGACTTTTCTCTAGAGCGCTCT 3180
QY 3181 GCGTTTTACTCTGTGTTTTCAAGGGACTTTTGTAAAATCAAAATCAATTCCTGTCAAGGCAG 3240
DB 3181 GCGTTTTACTCTGTGTTTTCAAGGGACTTTTGTAAAATCAAAATCAATTCCTGTCAAGGCAG 3240
QY 3241 GAGGAGTGTAAATGAACCTTTTATGGAGCATGATCTGCATCCAGGCGCTTCTCAGGCG 3300
DB 3241 GAGGAGTGTAAATGAACCTTTTATGGAGCATGATCTGCATCCAGGCGCTTCTCAGGCG 3300
QY 3301 GCTTGAGTGAATTTGTTACCTGAAGTACAGTATATCTTGTGAAAATACATAAACAAAGC 3360
DB 3301 GCTTGAGTGAATTTGTTACCTGAAGTACAGTATATCTTGTGAAAATACATAAACAAAGC 3360

Db 3301 GCTTGAGTGAATTGTACCTGAAGTACAGTATATCTTTGTAATACATAAAACAAAGC 3360
QY 3361 ATTTTCTAAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
Db 3361 ATTTTCTAAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
QY 3421 TTTTTCAGCTATTAGTGTATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
Db 3421 TTTTTCAGCTATTAGTGTATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 4
US-08-601-891-3
; Sequence 3, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,891
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..3039
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 139..3036
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..138
; US-08-601-891-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGGGCGCATCCGAGGGCTGGCGCGCGCCCTGGGGAGCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGGCGCATCCGAGGGCTGGCGCGCGCCCTGGGGAGCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGGCGTTGGCGGCGGACGCGTGGCGCTGCTGCTGTTGTTTTCATGCAATGATA 120
Db 61 CCGGCGTTGGCGGCGGACGCGTGGCGCTGCTGCTGTTGTTTTCATGCAATGATA 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
QY 181 AACAAATGATTCATCAGTGGGGAAGTCATCATCATCCATGTTATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTCATCAGTGGGGAAGTCATCATCATCCATGTTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAACTGCCCTGTG 300
Db 241 GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAACTGCCCTGTG 300
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGAACATT 360
Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGAACATT 360
QY 361 TCCTGTCTCTGGGCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTTACAA 420
Db 361 TCCTGTCTCTGGGCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTTACAA 420
QY 421 AACAGAGAGTGTGTTCCATGTCATTTTGAATGACAGAAACCCAGCTGAGAAATAC 480
Db 421 AACAGAGAGTGTGTTCCATGTCATTTTGAATGACAGAAACCCAGCTGAGAAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACACAATATTGTTTACAGTCAAGTAA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACACAATATTGTTTACAGTCAAGTAA 540
QY 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATCGAAAACAGGAGCGCC 600
Db 541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATCGAAAACAGGAGCGCC 600
QY 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGCTTGGCATTCAC 660
Db 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGCTTGGCATTCAC 660
QY 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720
Db 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720

QY 721 CATGAATATTGGAGCGACATAAGGTGCTGTGCCAGAAATGAACCTGGCAGGGAATGC 780
DB 721 CATGAATATTGGAGCGACATAAGGTGCTGTGCCAGAAATGAACCTGGCAGGGAATGC 780
QY 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAAATATTT 840
DB 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAAATATTT 840
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTCAATGTGAACCATGGATTC 900
DB 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTCAATGTGAACCATGGATTC 900
QY 901 GGGCTCACCTGGGAATAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960
DB 901 GGGCTCACCTGGGAATAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960
QY 961 ACCATTTCAACAACAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTGGCA 1020
DB 961 ACCATTTCAACAACAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTGGCA 1020
QY 1021 AGAAACGACACCGGATACACTGTTCTCCTTCAAGCATCCAGTCAATCAGCTTTG 1080
DB 1021 AGAAACGACACCGGATACACTGTTCTCCTTCAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCGTAGGAAGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140
DB 1081 GTTACCATCGTAGGAAGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140
QY 1141 GACCAATATGAAGAGTTTGTGTTTCTGTGAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
DB 1141 GACCAATATGAAGAGTTTGTGTTTCTGTGAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
QY 1201 ACCTGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACCGGATAC 1260
DB 1201 ACCTGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACCGGATAC 1260
QY 1261 AGCATATCCAACTTTGCAATCATAGCACACCCAGGAGAAATATATATTTCCATGCAGAA 1320
DB 1261 AGCATATCCAACTTTGCAATCATAGCACACCCAGGAGAAATATATATTTCCATGCAGAA 1320
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAAGAGAAACCTCAAGTG 1380
DB 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAAGAGAAACCTCAAGTG 1380
QY 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTCTGTTCTCGGATGGATACCATTAACCATCT 1440
DB 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTCTGTTCTCGGATGGATACCATTAACCATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATACAGAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATACAGAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGTCGACAGTACTCTA 1560
DB 1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGTCGACAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620
QY 1621 ACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCAGAGCAAC 1680
DB 1621 ACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCAGAGCAAC 1680
QY 1681 ATCTCATTTATGCAACAATTTGGTGTGCTCTCTTCAATTCGTTTTAAACCTGCTA 1740
DB 1681 ATCTCATTTATGCAACAATTTGGTGTGCTCTCTTCAATTCGTTTTAAACCTGCTA 1740
QY 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800

QY 1801 ACCGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
DB 1801 ACCGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
QY 1861 AATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
DB 1861 AATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGAAGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTTT 1980
DB 1921 GGAAGAAGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTTT 1980
QY 1981 GCCGTCAAAATGCTTGAAGAAAAGACAGACAGCTCTGAAAAGAGGCACTCATGTCAAAA 2040
DB 1981 GCCGTCAAAATGCTTGAAGAAAAGACAGACAGCTCTGAAAAGAGGCACTCATGTCAAAA 2040
QY 2041 CTCGAAGTATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTCTGGGGCGTGC 2100
DB 2041 CTCGAAGTATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTCTGGGGCGTGC 2100
QY 2101 ACATGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGATGTTGTTGATCTTCTCAAC 2160
DB 2101 ACATGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGATGTTGTTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAATTTACAGAGGACTTGCAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAAATTTACAGAGGACTTGCAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTCAAGTTTACCACCTTCCCAATCAATCAATCCAAATCCAGCATGCCCTGGTTCAAGA 2280
DB 2221 AATTTCAAGTTTACCACCTTCCCAATCAATCCAAATCCAGCATGCCCTGGTTCAAGA 2280
QY 2281 GAAGTTCAAGTATACACCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTCAC 2340
DB 2281 GAAGTTCAAGTATACACCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTCAC 2340
QY 2341 TCTGAAGATGAATGAATGATGAAGCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
DB 2341 TCTGAAGATGAATGAATGATGAAGCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGTTTGCATATCAAGTTGCAAGGAAATGGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTGTTTGCATATCAAGTTGCAAGGAAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGCGCCAGGAACGTCGTTTGTACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGCGCCAGGAACGTCGTTTGTACCCAC 2520
QY 2521 GGAAGAAGTGTGAAGATATGAGTGTGAGTTGGCTCGAGATATCATGAGTATTCACAC 2580
DB 2521 GGAAGAAGTGTGAAGATATGAGTGTGAGTTGGCTCGAGATATCATGAGTATTCACAC 2580
QY 2581 TATGTTCTCAGGGCAATGCCGCTGCTGCTGTAATTTGGATGGCCCGCCGAAAGCTGTTT 2640
DB 2581 TATGTTCTCAGGGCAATGCCGCTGCTGCTGTAATTTGGATGGCCCGCCGAAAGCTGTTT 2640
QY 2641 GAAGGATCTACACCATTAAGAGTGTGTCGTCATATGGAATATTTACTGTGSGAAATC 2700
DB 2641 GAAGGATCTACACCATTAAGAGTGTGTCGTCATATGGAATATTTACTGTGSGAAATC 2700
QY 2701 TTTCTCACTGTTGTAATCTTACCCTGGCATTCGGTGTGATGCTTAATCTTCAAACTG 2760
DB 2701 TTTCTCACTGTTGTAATCTTACCCTGGCATTCGGTGTGATGCTTAATCTTCAAACTG 2760
QY 2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAATAATACATATA 2820
DB 2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAATAATACATATA 2820
QY 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880
DB 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880
QY 2881 TTTTGTAGGATGTCAGCTGGCAGATGCGAAGAACCGGATGATCAGAATGTGGATGGCGGT 2940

Db 2881 TTTTGGATGTCAGTGGCAGATGCGAAGCGATGATCAGAATGTGATGCGCCT 2940
Qy 2941 GTTTCGGATGTCCTACACCTACCAACACAGCGGACCTTTTCAGCAGAGATGATTTG 3000
Db 2941 GTTTCGGATGTCCTACACCTACCAACACAGCGGACCTTTTCAGCAGAGATGATTTG 3000
Qy 3001 GGGCTACTCTCCCGAGGCTCAGTTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCCCGAGGCTCAGTTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAACAAAGATTAATTTAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAACAAAGATTAATTTAT 3120
Qy 3121 CACTAAAGAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAGCGCTCT 3180
Db 3121 CACTAAAGAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAGCGCTCT 3180
Qy 3181 GGGTTTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCAAGGCGAG 3240
Db 3181 GGGTTTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCAAGGCGAG 3240
Qy 3241 GAGGAGCTGATATGAATCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGG 3300
Db 3241 GAGGAGCTGATATGAATCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGG 3300
Qy 3301 GGTGAGTGAATGCTGACTGAAGTACAGATATATTCTTGTAAATCAATCATCTGTCAAGGCGG 3360
Db 3301 GGTGAGTGAATGCTGACTGAAGTACAGATATATTCTTGTAAATCAATCATCTGTCAAGGCGG 3360
Qy 3361 ATTTTCTAGGAGAGCTAAATGATTTTTTAAAGTCTATGTTTTTAAATCAATCATCTGTCAAGGCGG 3420
Db 3361 ATTTTCTAGGAGAGCTAAATGATTTTTTAAAGTCTATGTTTTTAAATCAATCATCTGTCAAGGCGG 3420
Qy 3421 TTTTTCAGCTATTAGTGTATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTAGTGTATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 5
US-09-021-324-3
Sequence 3, Application US/09021324
Patent No. 5912133
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
PRIOR APPLICATION DATA:

Query Match 100.0%; Score 3501; DB 2; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGAGCGGCATCCGAGGCGTGGCGCGGCGCTGGGGACCCCGGGTCCCGAGGCGCATG 60
Db 1 CGAGCGGCATCCGAGGCGTGGCGCGGCGCTGGGGACCCCGGGTCCCGAGGCGCATG 60
Qy 61 CCGCGGTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 CCGCGGTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTCTGCAATGATA 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTCTGCAATGATA 180
Qy 181 AACATGATTCATCAGTGGGAAGTCATCATCATATCCATGATCCAGATCCCGCGAA 240
Db 181 AACATGATTCATCAGTGGGAAGTCATCATCATATCCATGATCCAGATCCCGCGAA 240
Qy 241 GACCTCGGGTGTGGTTGAGAGCCCGAGAGCTCAGGAGAGCTGAGAGAGCTGCGCTGTG 300
Db 241 GACCTCGGGTGTGGTTGAGAGCCCGAGAGCTCAGGAGAGCTGAGAGAGCTGCGCTGTG 300

Db 241 GACCTCGGCTGTCGGTTGAGACCCCGAGAGCTCAGGGACAGTGACGAAGCTGCCGCTGTG 300
QY 301 GAAGTGGATGATCTGCTTCCATCACACTCAAGTCTGCTGATGCTGCCAGGGAACATT 360
Db 301 GAAGTGGATGATCTGCTTCCATCACACTCAAGTCTGCTGATGCTGCCAGGGAACATT 360
QY 361 TCCTGTCTCTGGTCTTTTAAGCACAGCTCCCTGAAATGCGCAGCACATTTTGAATTTACAA 420
Db 361 TCCTGTCTCTGGTCTTTTAAGCACAGCTCCCTGAAATGCGCAGCACATTTTGAATTTACAA 420
QY 421 AACAGAGGAGTTCTTCCATGCTCATTTTCAAAATGACAGAAACCCAAAGCTGGAGAATAC 480
Db 421 AACAGAGGAGTTCTTCCATGCTCATTTTCAAAATGACAGAAACCCAAAGCTGGAGAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAATTAAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAATTAAGA 540
QY 541 AATACCTCTCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAGGCC 600
Db 541 AATACCTCTCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAGGCC 600
QY 601 CTGGCTGCTATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660
Db 601 CTGGCTGCTATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660
QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCTT 720
Db 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCTT 720
QY 721 CATGAATTTTGGGACGGACATAAGTGTGTCGCCAGAAATGAACCTGGCAGGGAATGC 780
Db 721 CATGAATTTTGGGACGGACATAAGTGTGTCGCCAGAAATGAACCTGGCAGGGAATGC 780
QY 781 ACCAGGCTGTTCAATATAGATCTTAATCAAACTCCTCAGACCACATTTGCCAATTAATTT 840
Db 781 ACCAGGCTGTTCAATATAGATCTTAATCAAACTCCTCAGACCACATTTGCCAATTAATTT 840
QY 841 CTTAAAGTAGGGAACCCCTTATGGAATAGGTGCAAGCTGTTTCAATGTGAACCATGGATTC 900
Db 841 CTTAAAGTAGGGAACCCCTTATGGAATAGGTGCAAGCTGTTTCAATGTGAACCATGGATTC 900
QY 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCAGGAGGGCAACTACTTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCAGGAGGGCAACTACTTTTGAGATGAGT 960
QY 961 ACCTATTCACAAACAGACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020
Db 961 ACCTATTCACAAACAGACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020
QY 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCGATCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCGATCAATCAGCTTTG 1080
QY 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAATTT 1140
Db 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAATTT 1140
QY 1141 GACCAATATGAGAGTTTGTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
Db 1141 GACCAATATGAGAGTTTGTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
QY 1201 ACCTGACCTCTCTCGAAATCATTTCTTGTGACCAAAAGGCTTTGATAACGGATAC 1260
Db 1201 ACCTGACCTCTCTCGAAATCATTTCTTGTGACCAAAAGGCTTTGATAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTTCGAATATAAGCACAGCCAGGGAATATATATTTCCATGCAAGAA 1320
Db 1261 AGCATATCCAAGTTTTCGAATATAAGCACAGCCAGGGAATATATATTTCCATGCAAGAA 1320
QY 1321 AATGATGATGCCAAATTTACCAAAATGTTTCAACGCTGAATATAAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAAATTTACCAAAATGTTTCAACGCTGAATATAAGAGGAAACCTCAAGTG 1380

QY 1381 CTCGCAAGACATCGGCAAGTCAAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATCT 1440
Db 1381 CTCGCAAGACATCGGCAAGTCAAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATCT 1440
QY 1441 TGGACCTTGGAAAGTGTTCACAGCAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500
Db 1441 TGGACCTTGGAAAGTGTTCACAGCAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTCTCAGAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTCTCAGAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCATATAAAGGGTCTCCTGCTCAAGTCTGTGCATACAAATTCCTCTGC 1620
Db 1561 AACATGAGTGAAGCATATAAAGGGTCTCCTGCTCAAGTCTGTGCATACAAATTCCTCTGC 1620
QY 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680
Db 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTCTTCATTTGCTGTTTAAACCTTGTA 1740
Db 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTCCTCTCTTCATTTGCTGTTTAAACCTTGTA 1740
QY 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800
Db 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTTGATTTTCAGAGAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTTGATTTTCAGAGAATATGAATATGATCTC 1860
QY 1861 AAATGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGCTGATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAGCTGATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGACAGCTCTGAAAGAGAGGCACTCATGTGCAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAAGAAAAAGACAGCTCTGAAAGAGAGGCACTCATGTGCAGAA 2040
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100
QY 2101 ACAGTGTGAGGCAAAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160
Db 2101 ACAGTGTGAGGCAAAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160
QY 2161 TATCTAAGAAAGTAAAGAAAAATTTTACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAAAGTAAAGAAAAATTTTACAGGACTTTGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTAGTTTACCCTTTCATCATCCAAATTTCCAAATTTCCAGATGCTTCAAGA 2280
Db 2221 AATTTAGTTTACCCTTTCATCATCCAAATTTCCAAATTTCCAGATGCTTCAAGA 2280
QY 2281 GAAGTTTCAGATACACCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
Db 2281 GAAGTTTCAGATACACCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATTAATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATTAATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTCTTTTGTCTTATCAAGTTTGCACAAAGGAAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTCTTTTGTCTTATCAAGTTTGCACAAAGGAAATGGAATTT 2460

Qy	361	TCCTGCTCTCGGGTCTTTAAGCACAGCTCCCTGNAATTTGCCAGCCACATTTTGTATTACAA	420
Db	361		
Qy	361	TCCTGTCTCTGGGGTCTTTAAGCACAGCTCCCTGAAATTTGCCAGCCACATTTTGTATTACAA	420
Db	361		
Qy	421	AACAGAGGAGTGTGTTTCCATGGTTCATTTTGAAATGACAGAAACCCAAAGCTGGAGAAATAC	480
Db	421		
Qy	421	AACAGAGGAGTGTGTTTCCATGGTTCATTTTGAAATGACAGAAACCCAAAGCTGGAGAAATAC	480
Db	421		
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACAAATATTTTACAGTGAATTAAGA	540
Db	481		
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACAAATATTTTACAGTGAATTAAGA	540
Db	481		
Qy	541	AATACCCCTGCTTTACACATTTAAGAGACCTTACTTTAGAAAATTTGAAAACCCAGGAGGCC	600
Db	541		
Qy	541	AATACCCCTGCTTTACACATTTAAGAGACCTTACTTTAGAAAATTTGAAAACCCAGGAGGCC	600
Db	541		
Qy	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTCGGATTCA	660
Db	601		
Qy	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTCGGATTCA	660
Db	601		
Qy	661	CAGGGGAAAGCTGTTAAAGAAAGATCCAGCTGTGTTTAAAAGGAGGAAAAAGTGCCTT	720
Db	661		
Qy	661	CAGGGGAAAGCTGTTAAAGAAAGATCCAGCTGTGTTTAAAAGGAGGAAAAAGTGCCTT	720
Db	661		
Qy	721	CATGAATTTATTTGGGACGGACATAAGGTGCTGTGCCAGAAATGAATCGGCAGGGAAATGC	780
Db	721		
Qy	721	CATGAATTTATTTGGGACGGACATAAGGTGCTGTGCCAGAAATGAATCGGCAGGGAAATGC	780
Db	721		
Qy	781	ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAATGTCGCACAAATATTTT	840
Db	781		
Qy	781	ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAATGTCGCACAAATATTTT	840
Db	781		
Qy	841	CTTAAAGTAGGGAAACCTTTATGATAAGGTGCAAAAGCTGTCATGTGAACCATGGATTTC	900
Db	841		
Qy	841	CTTAAAGTAGGGAAACCTTTATGATAAGGTGCAAAAGCTGTCATGTGAACCATGGATTTC	900
Db	841		
Qy	901	GGGCTCACCTGGGAAATTTAGAAAACAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960
Db	901		
Qy	901	GGGCTCACCTGGGAAATTTAGAAAACAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960
Db	901		
Qy	961	ACCTATTCAAAACAAGCAACTATGATACGATCTGTGTTCTCTTTGTPATCATCAGTGGCA	1020
Db	961		
Qy	961	ACCTATTCAAAACAAGCAACTATGATACGATCTGTGTTCTCTTTGTPATCATCAGTGGCA	1020
Db	961		
Qy	1021	AGAAACGACACCCGGATCTACACTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG	1080
Db	1021		
Qy	1021	AGAAACGACACCCGGATCTACACTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG	1080
Db	1021		
Qy	1081	GTTTACCATCTGAGGAAGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT	1140
Db	1081		
Qy	1081	GTTTACCATCTGAGGAAGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT	1140
Db	1081		
Qy	1141	GACCAATATCAAGATTTTCTGTTCTGTCAGGTTTAAAGCCATCCACAAAATCAGATGT	1200
Db	1141		
Qy	1141	GACCAATATCAAGATTTTCTGTTCTGTCAGGTTTAAAGCCATCCACAAAATCAGATGT	1200
Db	1141		
Qy	1201	ACGTGGACCTTCTCTGAAAATCATTTCCCTGTGAGCAAAAGGCTCTTGATAACGGATAC	1260
Db	1201		
Qy	1201	ACGTGGACCTTCTCTGAAAATCATTTCCCTGTGAGCAAAAGGCTCTTGATAACGGATAC	1260
Db	1201		
Qy	1261	AGCATATCCAAGTTTGTGCAATCATAGCACCCAGGAGGAATATATATTTCCATGCGAGAA	1320
Db	1261		
Qy	1261	AGCATATCCAAGTTTGTGCAATCATAGCACCCAGGAGGAATATATATTTCCATGCGAGAA	1320
Db	1261		
Qy	1321	AATGATGATCCCAATTTTACCAAAATGTTTCCAGCTGAATATAGAGGAACCTCAAGTG	1380
Db	1321		
Qy	1321	AATGATGATCCCAATTTTACCAAAATGTTTCCAGCTGAATATAGAGGAACCTCAAGTG	1380
Db	1321		
Qy	1381	CTCGGAGAACATCGGCAAGTCAGGGCTGCTGTTTCTCGGATCGATACCATTTACCATCT	1440
Db	1381		
Qy	1381	CTCGGAGAACATCGGCAAGTCAGGGCTGCTGTTTCTCGGATCGATACCATTTACCATCT	1440
Db	1381		

QY	1441	TGGACCTGGGAAGAGTGTTCAGACAAAGTCTCCCACTGCGACAGAAAGAGATCAGAGAAGGA	1500
DB	1441	TGGACCTGGGAAGAGTGTTCAGACAAAGTCTCCCACTGCGACAGAAAGAGATCAGAGAAGGA	1500
QY	1501	GTCTGGAATAGAAAGGCTTAACAGAAAAGTCTTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
DB	1501	GTCTGGAATAGAAAGGCTTAACAGAAAAGTCTTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
QY	1561	AACATGAGTGAAGCCATAAAAGGTTCCCTGGTCAAAGTGTGTCATACAAATTCCTTGGC	1620
DB	1561	AACATGAGTGAAGCCATAAAAGGTTCCCTGGTCAAAGTGTGTCATACAAATTCCTTGGC	1620
QY	1621	ACATCTTTGTGAGAGATCCTTTTAAACTCTCAGAGCCCTTCCCTTTTCAATCCAAAGACAAC	1680
DB	1621	ACATCTTTGTGAGAGATCCTTTTAAACTCTCAGAGCCCTTCCCTTTTCAATCCAAAGACAAC	1680
QY	1681	ATCTCAATCTATGCAACAAATTTGGTTGTTGCTCCTCTTCAATGTCGTTTTTAAACCCCTGCTA	1740
DB	1681	ATCTCAATCTATGCAACAAATTTGGTTGTTGCTCCTCTTCAATGTCGTTTTTAAACCCCTGCTA	1740
QY	1741	ATTTGTCCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800
DB	1741	ATTTGTCCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800
QY	1801	ACGGCTCCTCAGATAATGAGTACTCTACGTTGATTCAGAGAAATATGAATATGATCTC	1860
DB	1801	ACGGCTCCTCAGATAATGAGTACTCTACGTTGATTCAGAGAAATATGAATATGATCTC	1860
QY	1861	AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920
DB	1861	AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920
QY	1921	GGAAAAGTGATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
DB	1921	GGAAAAGTGATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
QY	1981	CGCGTCAAATGTGAAAGAAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040
DB	1981	CGCGTCAAATGTGAAAGAAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040
QY	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATGTGAACCTGCTGGGGGGCGTGC	2100
DB	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATGTGAACCTGCTGGGGGGCGTGC	2100
QY	2101	ACACTGTCAAGACCAATTTACTTGATTTTGAATCTGTTGCTATGGTGTATCTTCTCAAC	2160
DB	2101	ACACTGTCAAGACCAATTTACTTGATTTTGAATCTGTTGCTATGGTGTATCTTCTCAAC	2160
QY	2161	TATCTAAGAGTAAAGAGAAAAATTTCAAGAGCTTGGACAGAGATTTTCAAGGAACAC	2220
DB	2161	TATCTAAGAGTAAAGAGAAAAATTTCAAGAGCTTGGACAGAGATTTTCAAGGAACAC	2220
QY	2221	AAATTCAGTTTTTACCCCACTTTTCCAATCACATCCAAATTCAGCATGCTGGTTCAGAA	2280
DB	2221	AAATTCAGTTTTTACCCCACTTTTCCAATCACATCCAAATTCAGCATGCTGGTTCAGAA	2280
QY	2281	GAGCTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATCGGGAATTCATTTCAAC	2340
DB	2281	GAGCTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATCGGGAATTCATTTCAAC	2340
QY	2341	TCTGAAGATGAAATTTGAATATGAAACCCAAAAGGCTGGAAGAGAGGAGGACTTTGAAT	2400
DB	2341	TCTGAAGATGAAATTTGAATATGAAACCCAAAAGGCTGGAAGAGAGGAGGACTTTGAAT	2400
QY	2401	GTGCTTACATTTGAAGATCTTCTTGGTTTGCATATCAAGTTGCCAAAAGATTTGAATTT	2460
DB	2401	GTGCTTACATTTGAAGATCTTCTTGGTTTGCATATCAAGTTGCCAAAAGATTTGAATTT	2460
QY	2461	CTGGAATTTTAAAGTCGTGTTACAGAGACCTTGGCCGCCAGGAACGTGCTGTCACCCAC	2520
DB	2461	CTGGAATTTTAAAGTCGTGTTACAGAGACCTTGGCCGCCAGGAACGTGCTGTCACCCAC	2520
QY	2521	GGAAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTTCCAAAC	2580

2521	DB	GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCAC	2580
2581	QY	TATGTTGTTCAGGGCAAATGCCCGTCTGCCTGTAAAAATGGATGGCCCCGAAAGCCGTGTTT	2640
2581	DB	TATGTTGTTCAGGGCAAATGCCCGTCTGCCTGTAAAAATGGATGGCCCCGAAAGCCGTGTTT	2640
2641	QY	GAAGGCATCTACACCATTAAGAGTGTCTGGTCATATGGGAATATTACTGTGGGAATC	2700
2641	DB	GAAGGCATCTACACCATTAAGAGTGTCTGGTCATATGGGAATATTACTGTGGGAATC	2700
2701	QY	TTTCTCACATGGTGTGAATCCTTTACCTTCGCATTCGGTTGATGCTAACTTCTACAAATG	2760
2701	DB	TTTCTCACATGGTGTGAATCCTTTACCTTCGCATTCGGTTGATGCTAACTTCTACAAATG	2760
2761	QY	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2761	DB	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2821	QY	ATGCAATCCTGCTGGGCTTTTGACTCAAGSAAAGGCCATCCTTCCTTAATTGACTTCG	2880
2821	DB	ATGCAATCCTGCTGGGCTTTTGACTCAAGSAAAGGCCATCCTTCCTTAATTGACTTCG	2880
2881	QY	TTTTTAGGATGTCAAGTGGCAGATGCAGAAGCGATGTATCAGAAATGTGGATGCCGT	2940
2881	DB	TTTTTAGGATGTCAAGTGGCAGATGCAGAAGCGATGTATCAGAAATGTGGATGCCGT	2940
2941	QY	GTTTCGGAATGCTCTCACACCTTACCAAAACAGCGGACCTTTCAGCAGAGATGGATTG	3000
2941	DB	GTTTCGGAATGCTCTCACACCTTACCAAAACAGCGGACCTTTCAGCAGAGATGGATTG	3000
3001	QY	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGG	3060
3001	DB	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGG	3060
3061	QY	ACTTCATCCCTCCACCTATGCCCTAACAGGCTGTAGATTTACCAAAACAAGATTAAATTCAT	3120
3061	DB	ACTTCATCCCTCCACCTATGCCCTAACAGGCTGTAGATTTACCAAAACAAGATTAAATTCAT	3120
3121	QY	CACATAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAAGCCGTCT	3180
3121	DB	CACATAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAAGCCGTCT	3180
3181	QY	CGCTTTACTCTGTTTTCAAAGGACCTTTTGTAAAAATCAATCATCTCTCAAGGCAG	3240
3181	DB	CGCTTTACTCTGTTTTCAAAGGACCTTTTGTAAAAATCAATCATCTCTCAAGGCAG	3240
3241	QY	GAGGAGCTGATATGAACCTTTATTGGAGCATTCGATCCCAAGGCCCTTCTCAGGCCG	3300
3241	DB	GAGGAGCTGATATGAACCTTTATTGGAGCATTCGATCCCAAGGCCCTTCTCAGGCCG	3300
3301	QY	GCTTGAGTGAATTTGTTACCTGAAATGACATATCTTCTTAATACATAAAACAAGC	3360
3301	DB	GCTTGAGTGAATTTGTTACCTGAAATGACATATCTTCTTAATACATAAAACAAGC	3360
3361	QY	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTAAATTAATATGAAA	3420
3361	DB	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTAAATTAATATGAAA	3420
3421	QY	TTTTTTCAGCTATTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA	3480
3421	DB	TTTTTTCAGCTATTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA	3480
3481	QY	AAAAAAAAAAAAAAAAAAAAA 3501	
3481	DB	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 7
PCT-US92-09893-3
; Sequence 3, Application PC/TUS9209893
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ImClone Systems Incorporated
 STREET: 180 Varick Street
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10014
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/09893
 FILING DATE: 19921116
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Feit, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM-3-7PT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3501 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 58..3039
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 139..3036
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 58..138
 PCT-US92-09893-3

	Query Match	100.0%;	Score 3501;	DB 5;	Length 3501;
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	Indels
	Matches 3501;	Conservative 0;	Gaps 0;		
Qy	1	CGAGGCGGCATCCGAGGGCTGGGGCGGCCCTGGGGACCCC	60		
Db	1	CGAGGCGGCATCCGAGGGCTGGGGCGGCCCTGGGGACCCC	60		
Qy	61	CCGCGGTGGGCGCGCAGCGGGCACCGTGCCTGCTCGTGTTGTTTTCTGCAATGATA	120		
Db	61	CCGCGGTGGGCGCGCAGCGGGCACCGTGCCTGCTCGTGTTGTTTTCTGCAATGATA	120		
Qy	121	TTTGGGACTATTACAATCAAGATCTGCGCTGTGATCAAGTGTGTTTTAATCAATCATAAAG	180		
Db	121	TTTGGGACTATTACAATCAAGATCTGCGCTGTGATCAAGTGTGTTTTAATCAATCATAAAG	180		
Qy	181	AACAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGCGAA	240		
Db	181	AACAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGCGAA	240		
Qy	241	GACCTCGGGTGTGGCTTGAGACCACAGAGCTCAGGCAGAGTGTACGAAGCTGCCGCTGTG	300		
Db	241	GACCTCGGGTGTGGCTTGAGACCACAGAGCTCAGGCAGAGTGTACGAAGCTGCCGCTGTG	300		
Qy	301	GAAGTGGATGTATCTGCCTTCCATCACAATGCAAGTGTGGTGCATGCCGCCAGGGAACATT	360		

Query Match	100.0%;	Score 3501;	DB 5;	Length 3501;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3501:	Conservative	0;	Mismatches	0;
			Indels	0;

Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTCTGTCGATGCCCCAGGGAACATT 360
QY 361 TCCTGTCTCTGGTCTTTTAAGCACAGCTCCCTGAATTGCCAGCCACATTTTGTATTTACAA 420
Db 361 TCCTGTCTCTGGTCTTTTAAGCACAGCTCCCTGAATTGCCAGCCACATTTTGTATTTACAA 420
QY 421 AACAGAGGAGTGTGTTCCTCATGTGTCATTTTGAATAATGACAGAAACCCCAAGCTGAGAAATAC 480
Db 421 AACAGAGGAGTGTGTTCCTCATGTGTCATTTTGAATAATGACAGAAACCCCAAGCTGAGAAATAC 480
QY 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATATACACAATATTTGTACAGTGAGTATAAGA 540
Db 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATATACACAATATTTGTACAGTGAGTATAAGA 540
QY 541 AATACCCCTCTTTTACACATTTAAGAGACCTTACTTTTAGAAAAATGGAATAACAGGAGCGC 600
Db 541 AATACCCCTCTTTTACACATTTAAGAGACCTTACTTTTAGAAAAATGGAATAACAGGAGCGC 600
QY 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCCGATTCA 660
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCCGATTCA 660
QY 661 CAGGGGGAAGCTGTAAGAAGAAAGTCAGCTGTTGTTTAAAGAGGAGAAAGTGCCT 720
Db 661 CAGGGGGAAGCTGTAAGAAGAAAGTCAGCTGTTGTTTAAAGAGGAGAAAGTGCCT 720
QY 721 CATGAATTTATTTGGGACCGACATAAGTCTGTGCCAGAAATGAACCTGGCAGGGAATGC 780
Db 721 CATGAATTTATTTGGGACCGACATAAGTCTGTGCCAGAAATGAACCTGGCAGGGAATGC 780
QY 781 ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTTATTT 840
Db 781 ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTTATTT 840
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGCAAAGCTGTTTCATGTGAACCATGCTTTC 900
Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGCAAAGCTGTTTCATGTGAACCATGCTTTC 900
QY 901 GGGCTCACCTGGGAATTAGAAACAAAGCACCTCGAGAGGGCACTACTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTAGAAACAAAGCACCTCGAGAGGGCACTACTTTGAGATGAGT 960
QY 961 ACCTATTCAACAAACAGAACTATGATACGATTTCTGTTTGTCTTGTATCATCATGCGCA 1020
Db 961 ACCATTCAACAAACAGAACTATGATACGATTTCTGTTTGTCTTGTATCATCATGCGCA 1020
QY 1021 AGAAGCACCGGATACACTTGTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAGCACCGGATACACTTGTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCTGAGGAAGGGATTTAATGCTACCAATTTCAAGTCAAGATTTGAAAT 1140
Db 1081 GTTACCATCTGAGGAAGGGATTTAATGCTACCAATTTCAAGTCAAGATTTGAAAT 1140
QY 1141 GACCAATATGAAGAGTTTTTGTCTGTCTGAGTAAAGCCTACCCACAAATCAGATGT 1200
Db 1141 GACCAATATGAAGAGTTTTTGTCTGTCTGAGTAAAGCCTACCCACAAATCAGATGT 1200
QY 1201 ACGTGGACCTTCTCTGGAATCATTTCTGTGTGAGCAAAAGGGCTTGTGATTAACGGATAC 1260
Db 1201 ACGTGGACCTTCTCTGGAATCATTTCTGTGTGAGCAAAAGGGCTTGTGATTAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACCCAGCCAGGAGGAATATATTTCCATGCAGAA 1320
Db 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACCCAGCCAGGAGGAATATATTTCCATGCAGAA 1320
QY 1321 AATGATGATGCCAATTTTACAAAATGTTTCAGCTGGAATATAAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTTACAAAATGTTTCAGCTGGAATATAAGAGGAAACCTCAAGTG 1380
QY 1381 CTCGAGAAAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGGATGGATACCCATTACCATCT 1440
Db 1381 CTCGAGAAAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGGATGGATACCCATTACCATCT 1440

QY 1441 TGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500
Db 1441 TGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTCCTCTCATTTAAACCTTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTCCTCTCATTTAAACCTTGGC 1620
QY 1621 ACATCTGTGAGAGTCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCCATCCAAAGAAC 1680
Db 1621 ACATCTGTGAGAGTCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCCATCCAAAGAAC 1680
QY 1681 ATCTCATTTTATGCAACAATTTGGTGTCTCTCTCTCATTTTAAACCTTGGT 1740
Db 1681 ATCTCATTTTATGCAACAATTTGGTGTCTCTCTCTCATTTTAAACCTTGGT 1740
QY 1741 ATTTGTCAAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800
Db 1741 ATTTGTCAAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800
QY 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
QY 1861 AAATGGAGTTTCCAAAGAGAAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGAGTTTCCAAAGAGAAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGTGTATGAAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAGTGTATGAAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY 1981 GCGCTCAAAATGCTGAAAGAAAGAGAGAGAGTCTGAAAGAGAGGAGTCTATGTCAGAA 2040
Db 1981 GCGCTCAAAATGCTGAAAGAAAGAGAGAGAGTCTGAAAGAGAGGAGTCTATGTCAGAA 2040
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTC 2100
QY 2101 ACATCTCAGGACCAATTTACTTTTGAATTTGTAATCTGTTGCTATGTTGCTCTCTCAAC 2160
Db 2101 ACATCTCAGGACCAATTTACTTTTGAATTTGTAATCTGTTGCTATGTTGCTCTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
Db 2221 AATTTTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTTCCAGCATGCTGTTCAAGA 2280
QY 2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
Db 2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAGAGAGGAGGAGTCTGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAGAGAGGAGGAGTCTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTTTGCATATCAAGTTGCCAAAGGAATTTTGAAT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTTTGCATATCAAGTTGCCAAAGGAATTTTGAAT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTTCAGAGAGCTTGGCCGCCAGGAACTGTTGTGCCACAC 2520
Db 2461 CTGGAATTTAAGTCTGTGTTTTCAGAGAGCTTGGCCGCCAGGAACTGTTGTGCCACAC 2520

QY 421 AACAGAGAGTGTGTTCCATGGTCATTTGAAAAATGACAGAAACCCCAAGCTGGAGAATAC 480
DB 421 AACAGAGAGTGTGTTCCATGGTCATTTGAAAAATGACAGAAACCCCAAGCTGGAGAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACAAATATTTGTTTACAGTGAGTATAGA 540
DB 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACAAATATTTGTTTACAGTGAGTATAGA 540
QY 541 AATACCCCTGCTTACACATTAAGNAGACCTTACTTTAGAAAAATGGAACCCAGGAGCC 600
DB 541 AATACCCCTGCTTACACATTAAGNAGACCTTACTTTAGAAAAATGGAACCCAGGAGCC 600
QY 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGCGTCTTGGGATTCA 660
DB 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGCGTCTTGGGATTCA 660
QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
DB 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
QY 721 CATGAATTTTGGGAGGACATAAGGTGCTGTCGCGAGAAATGAACHTGGGAGGAATGC 780
DB 721 CATGAATTTTGGGAGGACATAAGGTGCTGTCGCGAGAAATGAACHTGGGAGGAATGC 780
QY 781 ACCAGGCTGTTCACATAGATCTAAATCAAACCTCTCAGACCACATTCGCCACATTTATTT 840
DB 781 ACCAGGCTGTTCACATAGATCTAAATCAAACCTCTCAGACCACATTCGCCACATTTATTT 840
QY 841 CTTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGATTC 900
DB 841 CTTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGATTC 900
QY 901 GGCTCACCTGGAAATTAGAAACAAAGCACTGAGAGGGCAACTACTTTGAGATGAGT 960
DB 901 GGCTCACCTGGAAATTAGAAACAAAGCACTGAGAGGGCAACTACTTTGAGATGAGT 960
QY 961 ACCTATTCAACAACAGCACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGCSCA 1020
DB 961 ACCTATTCAACAACAGCACTATGATACGGATCTGTTGCTTTTGTATCATCAGTGCSCA 1020
QY 1021 AGAACGACACCGGATCTACTACTTGTCTCTTCAAGGATCCCAAGTCAATCAGCTTTG 1080
DB 1021 AGAACGACACCGGATCTACTACTTGTCTCTTCAAGGATCCCAAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCTAGAAAGGATTTATATAATGCTTACCAATTCAGTGAAGATTTAGAAAT 1140
DB 1081 GTTACCATCTAGAAAGGATTTATATAATGCTTACCAATTCAGTGAAGATTTAGAAAT 1140
QY 1141 GACCAATATCAAGAGTTTCTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
DB 1141 GACCAATATCAAGAGTTTCTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
QY 1201 ACCTGGACCTTCTCTGAAATCATTTCTTGTGAGCAAAAGGCTCTGTATACGGATAC 1260
DB 1201 ACCTGGACCTTCTCTGAAATCATTTCTTGTGAGCAAAAGGCTCTGTATACGGATAC 1260
QY 1261 AGCATATCCAAAGTTTGCATATCATAAGCACACCGAGGAGAAATATATATTCATTCAGAA 1320
DB 1261 AGCATATCCAAAGTTTGCATATCATAAGCACACCGAGGAGAAATATATATTCATTCAGAA 1320
QY 1321 AATGATGATCCCAATTTTACCAAAATGTTTCAACGCTGAATATAGAAGGAAACCTCAAGTG 1380
DB 1321 AATGATGATCCCAATTTTACCAAAATGTTTCAACGCTGAATATAGAAGGAAACCTCAAGTG 1380
QY 1381 CTGCGAGAAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTAACCATCT 1440
DB 1381 CTGCGAGAAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTAACCATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCGACAGAAGAGATCACAGAAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCGACAGAAGAGATCACAGAAGGA 1500

QY 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAAGTGTGACAGTGGTGTGAGCAGTACTCTA 1560
DB 1501 GTCTGGAAATAGAAAGGCTTAACAGAAAAAGTGTGACAGTGGTGTGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAAGGGTCTCTGCTCAAGTCTGTGCATACAATTCCTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAAGGGTCTCTGCTCAAGTCTGTGCATACAATTCCTTGGC 1620
QY 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGCAAC 1680
DB 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGCAAC 1680
QY 1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAGGTG 1800
DB 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAGGTG 1800
QY 1801 ACCGCTCTCAGATATAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATGATCTC 1860
DB 1801 ACCGCTCTCAGATATAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATGATCTC 1860
QY 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTCTTTT 1920
DB 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTCTTTT 1920
QY 1921 GGAAAGTGTAGAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
DB 1921 GGAAAGTGTAGAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
QY 1981 GCCGTCAAAATGCTGAAGAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
DB 1981 GCCGTCAAAATGCTGAAGAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
QY 2041 CTCAGATGATGACCCAGCTGGAGGCCACAGAAATTTGTCAACCTCTGGGGCGTGC 2100
DB 2041 CTCAGATGATGACCCAGCTGGAGGCCACAGAAATTTGTCAACCTCTGGGGCGTGC 2100
QY 2101 ACATGTCAAGACCAATTTACTTGAATTTTGAATCTGCTATGTGATCTTCTCAAC 2160
DB 2101 ACATGTCAAGACCAATTTACTTGAATTTTGAATCTGCTATGTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTCAAGTTTACCCTTCCCAATCAGATCCAAATTCAGCATGCCCTGGTTCAGA 2280
DB 2221 AATTTCAAGTTTACCCTTCCCAATCAGATCCAAATTCAGCATGCCCTGGTTCAGA 2280
QY 2281 GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTCAC 2340
DB 2281 GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAAATGAAATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
DB 2341 TCTGAAGATGAAATGAAATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTTCATATCAAGTTTGCAGAAAGGATGGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTTCATATCAAGTTTGCAGAAAGGATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCAGAGACCTGGCCGAGGAGGAGTGTGTGCACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCAGAGACCTGGCCGAGGAGGAGTGTGTGCACCCAC 2520
QY 2521 GGGAAAGTGTGAAAGATATGTGACTTTTGGATTTGCTCGAGATATCATGAGTATTCACAC 2580
DB 2521 GGGAAAGTGTGAAAGATATGTGACTTTTGGATTTGCTCGAGATATCATGAGTATTCACAC 2580
QY 2581 TATCTTCTCAGGGGCAATGCCCTGTCTGCTGTAATAATGGATGGCCCCCGAAGGCTGTTT 2640

Db 2701 TTCTCAGTGTGTGAATCCTTACCTCGCATTCGCGTTGATGCTAACTTCTACAACTG 2760
QY 2761 ATTCAAAATGGATTTAAATAGATCAGCCATTTTATGCTACAGAAATATACATTATA 2820
Db 2761 ATTCAAAATGGATTTAAATAGATCAGCCATTTTATGCTACAGAAATATACATTATA 2820
QY 2821 ATGCAATCCTGCTGGCTTTTTCACCTCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCCTGCTGGCTTTTTCACCTCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
QY 2881 TTTTATGAGATGTCAGTGCAGATGCGAAGAGCGGATGATGATCAGAAATGCGATGCGCGT 2940
Db 2881 TTTTATGAGATGTCAGTGCAGATGCGAAGAGCGGATGATGATCAGAAATGCGATGCGCGT 2940
QY 2941 GTTTCGGAAATGCTTCCACACTACCAAAACAGCGGACCTTTTCAGCAGAGATGGATTG 3000
Db 2941 GTTTCGGAAATGCTTCCACACTACCAAAACAGCGGACCTTTTCAGCAGAGATGGATTG 3000
QY 3001 GGCTACTCTCCGCGAGCTCAGGTGCGAAGATTCTAGAGGAACAATTTAGTTTAAAG 3060
Db 3001 GGCTACTCTCCGCGAGCTCAGGTGCGAAGATTCTAGAGGAACAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGATGATTTACCAAAACAAGATTAAATTCAT 3120
Db 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGATGATTTACCAAAACAAGATTAAATTCAT 3120
QY 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTCCAGCACTTTTCTAGAGAGCGCTCT 3180
Db 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTCCAGCACTTTTCTAGAGAGCGCTCT 3180
QY 3181 GCGTTTACTCTCTTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCACAAAGCAG 3240
Db 3181 GCGTTTACTCTCTTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCACAAAGCAG 3240
QY 3241 GAGGAGCTGATATGAACCTTTATTGGAGCAATGATGTCATCAAGGCTTCTCAGGCGG 3300
Db 3241 GAGGAGCTGATATGAACCTTTATTGGAGCAATGATGTCATCAAGGCTTCTCAGGCGG 3300
QY 3301 GCTTGTAGTGAATGTGTACTGAGTACAGTATATCTTCTTAATACATAAACAAGC 3360
Db 3301 GCTTGTAGTGAATGTGTACTGAGTACAGTATATCTTCTTAATACATAAACAAGC 3360
QY 3361 ATTTTGTAGGAGAGCACTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
Db 3361 ATTTTGTAGGAGAGCACTAATATGATTTTAAAGTCTATGTTTAAATAATATGTA 3420
QY 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475
Db 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475

RESULT 10
PCT-US95-03718-3
; Sequence 3, Application PC/TUS9503718
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM: PCT/US95/03718
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 879PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3475 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-03718-3

Query Match 99.3%; Score 3475; DB 5; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGCGGATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGCGGATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGGCTTGGCGCGGACGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 120
Db 61 CCGGCTTGGCGCGGACGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTATCAAGTGTGTTTAAATCAATCAATAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTATCAAGTGTGTTTAAATCAATCAATAAG 180
QY 181 AACAAATGATTCATCAGTGGGGAAGTATCATATATCCATGGTATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTCATCAGTGGGGAAGTATCATATATCCATGGTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGGTGTGCGTTGAGACCCGAGAGCTCAGGACAGTGTACGAGCTGCGCTGTG 300
Db 241 GACCTCGGGTGTGCGTTGAGACCCGAGAGCTCAGGACAGTGTACGAGCTGCGCTGTG 300
QY 301 GAAGTGTATGATCTGCTTCCATCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GAAGTGTATGATCTGCTTCCATCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 TCCTGTCTCTGGTCTTTTAAAGCAGCTCCCTGAAATGCGGACGACATTTTGTATTACAA 420
Db 361 TCCTGTCTCTGGTCTTTTAAAGCAGCTCCCTGAAATGCGGACGACATTTTGTATTACAA 420
QY 421 AACAGAGAGTGTGTTTCCATGGTCTATTTGAAATGACAGAAACCCAGCTGGAGATAC 480
Db 421 AACAGAGAGTGTGTTTCCATGGTCTATTTGAAATGACAGAAACCCAGCTGGAGATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTTACAGTGTATTAAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTTACAGTGTATTAAGA 540
QY 541 AATACCTGCTTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCGCC 600
Db 541 AATACCTGCTTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCGCC 600
QY 601 CTGTCTCATATCTGAGAGGCTTCCAGAGCGGATCGTGGATGGTCTTGGCATTC 660
Db 601 CTGTCTCATATCTGAGAGGCTTCCAGAGCGGATCGTGGATGGTCTTGGCATTC 660
QY 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTTAAAGAGAGAAAAAGTGTCT 720
Db 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTTAAAGAGAGAAAAAGTGTCT 720
QY 721 CATGAATTTTGGGACGCGACATAGTGTCTGCTGCCAGAAATGAATGGCGGAGGAATGC 780

|||||
721 CATGAATATTGCGGACGACATAAGGTGCTGTGCCAGAAATGAACCTGGCGAGGAATGC 780
QY
781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCACATGTCACAAATATT 840
Db
781 ACCAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACCACATGTCACAAATATT 840
QY
841 CTTAAAGTAGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
Db
841 CTTAAAGTAGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
QY
901 GGGCTCACCTGGGAATTAGAAAACAAGCACTCGAGSAGGCAACTACTTTTGATGAGT 960
Db
901 GGGCTCACCTGGGAATTAGAAAACAAGCACTCGAGSAGGCAACTACTTTTGATGAGT 960
QY
961 ACCTATTCAACAACAGAACTATGATACGGATTCTGTTTCTGTTTGTATCATCAGTGGCA 1020
Db
961 ACCTATTCAACAACAGAACTATGATACGGATTCTGTTTCTGTTTGTATCATCAGTGGCA 1020
QY
1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCGAGTCAATCAGCTTTG 1080
Db
1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCGAGTCAATCAGCTTTG 1080
QY
1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTCAAAGTGAAGATTTATGAAT 1140
Db
1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTCAAAGTGAAGATTTATGAAT 1140
QY
1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200
Db
1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200
QY
1201 ACGTGACCTCTCTCGAAATCATTTCTGTTGAGCAAAAGGCTCTTGATAACGGATAC 1260
Db
1201 ACGTGACCTCTCTCGAAATCATTTCTGTTGAGCAAAAGGCTCTTGATAACGGATAC 1260
QY
1261 AGCATATCCAAGTTTGCATCATAGCACACGACGAGGAATATATATTCATTCAGAA 1320
Db
1261 AGCATATCCAAGTTTGCATCATAGCACACGACGAGGAATATATATTCATTCAGAA 1320
QY
1321 AATGATGATGCCAATTTACCAAAATGTTCAACCTGAAATAGAGAGAAACCTCAAGTG 1380
Db
1321 AATGATGATGCCAATTTACCAAAATGTTCAACCTGAAATAGAGAGAAACCTCAAGTG 1380
QY
1381 CTCGAGAAGCATCGCAAGTCAAGGCTCCTGTTCTCGGATGATACCATACCATCT 1440
Db
1381 CTCGAGAAGCATCGCAAGTCAAGGCTCCTGTTCTCGGATGATACCATACCATCT 1440
QY
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCACTGCACAGAAAGATCACAGAAGGA 1500
Db
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCACTGCACAGAAAGATCACAGAAGGA 1500
QY
1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGCGAGCAGTACTCTA 1560
Db
1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGCGAGCAGTACTCTA 1560
QY
1561 AACATGAGTGAAGCCATAAAGGTTCTGTTGTCAGTCAAGTGTGTCATACAAATCCCTTGGC 1620
Db
1561 AACATGAGTGAAGCCATAAAGGTTCTGTTGTCAGTCAAGTGTGTCATACAAATCCCTTGGC 1620
QY
1621 ACATCTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGCAAC 1680
Db
1621 ACATCTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGCAAC 1680
QY
1681 ATCTCATTTATGCAACAATGGTGTGTTCTCCTCTTCATTTGCTGTTTAAACCCCTGCTA 1740
Db
1681 ATCTCATTTATGCAACAATGGTGTGTTCTCCTCTTCATTTGCTGTTTAAACCCCTGCTA 1740
QY
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGTACAGATGGTACAGGTG 1800
Db
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGTACAGATGGTACAGGTG 1800
QY
1801 ACCGGCTCCTCAGATAATCAGTACTCTACGTGTGATTTTCAGAGAAATATGAATGATCTC 1860
|||||

Db
1801 ACCGGCTCCTCAGATAAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
QY
1861 AATGGGAGTTTCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
Db
1861 AATGGGAGTTTCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920
QY
1921 GGAAGTGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db
1921 GGAAGTGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY
1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACTCTGAAAGAGAGGCACTCATGTGAGAA 2040
Db
1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACTCTGAAAGAGAGGCACTCATGTGAGAA 2040
QY
2041 CTCAGATGATGACCCAGCTGGGAGCCACAGATAATTTGTAACCTGCTGGGGCGTGC 2100
Db
2041 CTCAGATGATGACCCAGCTGGGAGCCACAGATAATTTGTAACCTGCTGGGGCGTGC 2100
QY
2101 ACCTGTCCAGACCAATTTACTTCAATTTCAATCTGCTGATGTTGATCTTCTCAAC 2160
Db
2101 ACCTGTCCAGACCAATTTACTTCAATTTCAATCTGCTGATGTTGATCTTCTCAAC 2160
QY
2161 TATCTAAGAAAGTAAAAAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAC 2220
Db
2161 TATCTAAGAAAGTAAAAAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAC 2220
QY
2221 AATTTAGTTTTTACCCTACTTTCCAAATCATCCAAATCCAGATGCCCTGGTTCAAGA 2280
Db
2221 AATTTAGTTTTTACCCTACTTTCCAAATCATCCAAATCCAGATGCCCTGGTTCAAGA 2280
QY
2281 GAAGTTCCAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTTCC 2340
Db
2281 GAAGTTCCAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTTCC 2340
QY
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTTGAAGAGAGAGGACTTTGAAT 2400
Db
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTTGAAGAGAGAGGACTTTGAAT 2400
QY
2401 GTGTTACATTTGAAGATCTCTTTGCTTTCATATCAAGTTGCAAGAGAAATGGAATTT 2460
Db
2401 GTGTTACATTTGAAGATCTCTTTGCTTTCATATCAAGTTGCAAGAGAAATGGAATTT 2460
QY
2461 CTGGAATTTAAGTCGTGTTTCAGAGACCTGGCCGACAGAGCTGCTTGTCCACCCAC 2520
Db
2461 CTGGAATTTAAGTCGTGTTTCAGAGACCTGGCCGACAGAGCTGCTTGTCCACCCAC 2520
QY
2521 GGAAGTGTGAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTGTATCCAAAC 2580
Db
2521 GGAAGTGTGAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTGTATCCAAAC 2580
QY
2581 TATGTTCTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCTGTTT 2640
Db
2581 TATGTTCTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCTGTTT 2640
QY
2641 GAAGGCATCTACACCATTAAGAGTGTCTGTTGATATGGAATATTACTGTGGGAATC 2700
Db
2641 GAAGGCATCTACACCATTAAGAGTGTCTGTTGATATGGAATATTACTGTGGGAATC 2700
QY
2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTGTGATGCTAACTTCTCAAACTG 2760
Db
2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTGTGATGCTAACTTCTCAAACTG 2760
QY
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
Db
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
QY
2821 ATGCATCTGCTGGGCTTTTGTACTCAAGGAACGGCCATCTTCCCTTAATTTGACTTCG 2880
Db
2821 ATGCATCTGCTGGGCTTTTGTACTCAAGGAACGGCCATCTTCCCTTAATTTGACTTCG 2880
QY
2881 TTTTATAGGATGTCAGTGGCAGATGCAGAGAACCGGATGTATCAGATGTGGATGGCGGT 2940
Db
2881 TTTTATAGGATGTCAGTGGCAGATGCAGAGAACCGGATGTATCAGATGTGGATGGCGGT 2940

Qy	2941	GTTCGGAAATGCTCTCACACCTACCAAAACAGCGCACCTTTTCAGCAGAGAGATGGATTG	3000
Db	2941	GTTCGGAAATGCTCTCACACCTACCAAAACAGCGCACCTTTTCAGCAGAGAGATGGATTG	3000
Qy	3001	GGGCTACTCTCTCGGAGGCTCAGGTCGAAGATTCGTAGAGGAACAAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCTCGGAGGCTCAGGTCGAAGATTCGTAGAGGAACAAATTTAGTTTTAAGG	3060
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGCGCTAGATTACCAAAACAAGATTAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGCGCTAGATTACCAAAACAAGATTAATTTTCAT	3120
Qy	3121	CACATAAAGAAAAATCTAATTAACAACCTGCTGCTTCCACGAGACTTTTCTCTAGAAGCGCTCT	3180
Db	3121	CACATAAAGAAAAATCTAATTAACAACCTGCTGCTTCCACGAGACTTTTCTCTAGAAGCGCTCT	3180
Qy	3181	GCCTTTTACTCTCTTTTCAAGGAGCATTTTGTAAAAATCAATCATCCTGTGCACAAAGGCAG	3240
Db	3181	GCCTTTTACTCTCTTTTCAAGGAGCATTTTGTAAAAATCAATCATCCTGTGCACAAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACATTTATTGGAGCATTTGATCTGCATCCAAAGSCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACATTTATTGGAGCATTTGATCTGCATCCAAAGSCCTTCTCAGGCCG	3300
Qy	3301	GCTTGAAGTGAATTCGTGAACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAGAAGC	3360
Db	3301	GCTTGAAGTGAATTCGTGAACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAGAAGC	3360
Qy	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTTTAAAGTCTATCTTTTAAAAATAATATGTAAA	3420
Db	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTTTAAAGTCTATCTTTTAAAAATAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGTGGGAATAAAAATTTCTACTACAG	3475
Db	3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGTGGGAATAAAAATTTCTACTACAG	3475

RESULT 11

RESOLV 11
US-08-183-211-1
; Sequence 1, Application US/08183211
; Patent No. 5618709
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR
; TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STK-1 PROTEIN
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,211
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383

QY 901 GGGCTCACCTGGGAATTTAGAAAACAAACGACACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
DB 901 GGGCTCACCTGGGAATTTAGAAAACAAACGACACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
QY 961 ACCTATTTCAACAACAGAACTATGATACGGAATCTGTTGCTTTTGTATCATCATGAGTGCA 1020
DB 961 ACCTATTTCAACAACAGAACTATGATACGGAATCTGTTGCTTTTGTATCATCATGAGTGCA 1020
QY 1021 AGAAACGACACCGGATACACTTTGCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
DB 1021 AGAAACGACACCGGATACACTTTGCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAT 1140
DB 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAT 1140
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
DB 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
QY 1201 ACCTGACCTTCTCTGAAATCATTTCTTGTGTCAGCAAAAGGCTTTGATAACGGATAC 1260
DB 1201 ACCTGACCTTCTCTGAAATCATTTCTTGTGTCAGCAAAAGGCTTTGATAACGGATAC 1260
QY 1261 AGCATATCCAAATTTTGCATATCAAGCACGACGAGGAATATATATTCCATGCGAGAA 1320
DB 1261 AGCATATCCAAATTTTGCATATCAAGCACGACGAGGAATATATATTCCATGCGAGAA 1320
QY 1321 AATGATGATGCCAAATTTACCAAAATGTTCCAGCTGAATATAGAAGGAAACCTTCAAGTG 1380
DB 1321 AATGATGATGCCAAATTTACCAAAATGTTCCAGCTGAATATAGAAGGAAACCTTCAAGTG 1380
QY 1381 CTGCGAGAACATCGCAAGTCTAGGCTCTGTTTCTGATGATGATACCATTTACCATCT 1440
DB 1381 CTGCGAGAACATCGCAAGTCTAGGCTCTGTTTCTGATGATGATACCATTTACCATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGACAGAGAGATCACAGAAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGACAGAGAGATCACAGAAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTCAGAGTGGGTGTCGACAGTACTCTA 1560
DB 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTCAGAGTGGGTGTCGACAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTCTCGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620
DB 1561 AACATGAGTGAAGCCATAAAGGGTCTCGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620
QY 1621 ACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
DB 1621 ACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
QY 1681 ATCTCATTTATGCAACAATTTGTTGTTCTCTCTTCATTTGTCGTTTAAACCTGCTA 1740
DB 1681 ATCTCATTTATGCAACAATTTGTTGTTCTCTCTTCATTTGTCGTTTAAACCTGCTA 1740
QY 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATGATATGATCTC 1860
DB 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATGATATGATCTC 1860
QY 1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAAAGTACTAGGATCAGGTGCTTTT 1920
DB 1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAAAGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGATGATGAACCCACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
DB 1921 GGAAGATGATGAACCCACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980

QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAGAA 2040
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAGAA 2040
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
DB 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
QY 2101 ACACGTGTCAGACCAATTTACTTGAATTTTGAATACTGTTGATGTTGATGTTCTTCAC 2160
DB 2101 ACACGTGTCAGACCAATTTACTTGAATTTTGAATACTGTTGATGTTGATGTTCTTCAC 2160
QY 2161 TATCTAAGAACTAAAGAGAAAAATTTTCACAGGACTTGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAACTAAAGAGAAAAATTTTCACAGGACTTGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTCAAGTTTTTACCCCACTTTCCAAATCACATCCAAATTCAGCATGCTGGTTCAGA 2280
DB 2221 AATTTCAAGTTTTTACCCCACTTTCCAAATCACATCCAAATTCAGCATGCTGGTTCAGA 2280
QY 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
DB 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGA 2400
DB 2341 TCTGAAGATGAAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGA 2400
QY 2401 GTGCTTACATTTGAAGATCTTTTGTCTTTCATCAATCAAGTTCGCAAAAGGAATGGAAT 2460
DB 2401 GTGCTTACATTTGAAGATCTTTTGTCTTTCATCAATCAAGTTCGCAAAAGGAATGGAAT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGAGAACTGCTTGTTCACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGAGAACTGCTTGTTCACCCAC 2520
QY 2521 GGGAAAGTGTGAAGATATGAGTCTTGGATTTGCTCGAGATATCATGAGTATTCACAC 2580
DB 2521 GGGAAAGTGTGAAGATATGAGTCTTGGATTTGCTCGAGATATCATGAGTATTCACAC 2580
QY 2581 TATCTTGTGAGGGCAATGCCCCTGCTGCTGTAAATGGATGGCCCGGAAAGCCCTGTT 2640
DB 2581 TATCTTGTGAGGGCAATGCCCCTGCTGCTGTAAATGGATGGCCCGGAAAGCCCTGTT 2640
QY 2641 GAAGGCATCTACACCATTAAGAGTATGCTGTCATATGGAATATTTACTGTGGAAATC 2700
DB 2641 GAAGGCATCTACACCATTAAGAGTATGCTGTCATATGGAATATTTACTGTGGAAATC 2700
QY 2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTGTGATGCTTAACCTTCTACAACTG 2760
DB 2701 TTCTCACTTGGTGTGAATCCTTACCCCTGGCATTCGGTGTGATGCTTAACCTTCTACAACTG 2760
QY 2761 ATTCAAAATGGATTTAAATGGATCAGCAATTTTATGCTACAGAGAAATATACATATA 2820
DB 2761 ATTCAAAATGGATTTAAATGGATCAGCAATTTTATGCTACAGAGAAATATACATATA 2820
QY 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880
DB 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880
QY 2881 TTTTGTAGATGTCAGCTGGGAGATGCGAGAAAGCGATGATCAGATGTCGATGGCCGT 2940
DB 2881 TTTTGTAGATGTCAGCTGGGAGATGCGAGAAAGCGATGATCAGATGTCGATGGCCGT 2940
QY 2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCACAGAGAGATGGATTTG 3000
DB 2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCACAGAGAGATGGATTTG 3000
QY 3001 GGGCTACTCTCCGAGGCTCAGTCTGAGGATTTCTAGAGGAACAAATTTAGTTTAAAG 3060
DB 3001 GGGCTACTCTCCGAGGCTCAGTCTGAGGATTTCTAGAGGAACAAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGATAGATTACCAAAACAGAGATTAATTTTCAT 3120

3061	ACTTTCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTCAT	3120
3121	CACATAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAGCCGTCT	3180
3121	CACATAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAGCCGTCT	3180
3181	CGCTTTACCTCTGTTTTCAAAGGACTTTTGTAATAATCAAAATCATCTGTCACAAAGCAG	3240
3181	CGCTTTACCTCTGTTTTCAAAGGACTTTTGTAATAATCAAAATCATCTGTCACAAAGCAG	3240
3241	GAGGAGCTGATATGAATGAATTTATTGGAGCAATTCATCTGCATCCAAGGCCCTTCAGGCCG	3300
3241	GAGGAGCTGATATGAATGAATTTATTGGAGCAATTCATCTGCATCCAAGGCCCTTCAGGCCG	3300
3301	GCPTTGAGTGAATTTGCTACCTGGAATACAGTATATCTTGTAAATACATAAAACAAGAAGC	3360
3301	GCPTTGAGTGAATTTGCTACCTGGAATACAGTATATCTTGTAAATACATAAAACAAGAAGC	3360
3361	ATTTTTCGTAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTTAAAAATAATGTAAA	3420
3361	ATTTTTCGTAGGAGAAGCTAATATGATTTTTTAAGTCTATGTTTTTAAAAATAATGTAAA	3420
3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGA	3476
3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGA	3476

RESULT 12
 PCT-US95-00176A-1
 ; Sequence 1, Application PC/TUS9500176A
 ; GENERAL INFORMATION:
 ; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
 ; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
 ; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR INHIBITING EXPRESSION OF THE
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
 ; STREET: Suite 1800, Penn Center Plaza
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/00176A
 ; FILING DATE: 6 January 1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/183,211
 ; FILING DATE: 14 January 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monaco, Daniel A.
 ; REGISTRATION NUMBER: 30,480
 ; REFERENCE/DOCKET NUMBER: 3957-14 PC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-8383
 ; TELEFAX: (215) 568-5549
 ; TELEX: None
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3476 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single stranded
 ; TOPOLOGY: linear
 ; PCT-US95-00176A-1

Query Match 99.2%; Score 3474.4; DB 5; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;

Matches	3475:	Conservative	0:	Mismatches	1:	Indels	Q:	Gaps	0:
Qy	1	CGAGGGCGCATCGAGGGCTGGCGCGGCCCTGGGGGACCCGGGCTCGGAGGCCATG	60						
Db	1	CGAGGGCGCATCGAGGGCTGGCGCGGCCCTGGGGGACCCGGGCTCGGAGGCCATG	60						
Qy	61	CCGGCGTTGGCGGCGGAGCGGGGACCGTGCCGCTGCTCGTTGTTTTCTGCAATGATA	120						
Db	61	CCGGCGTTGGCGGCGGAGCGGGGACCGTGCCGCTGCTCGTTGTTTTCTGCAATGATA	120						
Qy	121	TTTGGGACTATTACAAATCAAGATCTCCCTGTGATCAAGTGCTGTTTAAATCAATCATAAAG	180						
Db	121	TTTGGGACTATTACAAATCAAGATCTCCCTGTGATCAAGTGCTGTTTAAATCAATCATAAAG	180						
Qy	181	AACAAATGATTATCATAGTGGGGAAGTCATCATATATCCCATGGTATCAGAAATCCCCGGAA	240						
Db	181	AACAAATGATTATCATAGTGGGGAAGTCATCATATATCCCATGGTATCAGAAATCCCCGGAA	240						
Qy	241	GACCTCGGTTGGGTTGAGACCCAGACTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300						
Db	241	GACCTCGGTTGGGTTGAGACCCAGACTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300						
Qy	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGCTGGTCCGATGCCCGAGGGAACATT	360						
Db	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGCTGGTCCGATGCCCGAGGGAACATT	360						
Qy	361	TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTGCCAGCCACATTTTGATTTACAA	420						
Db	361	TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTGCCAGCCACATTTTGATTTACAA	420						
Qy	421	AACAGAGAGTGTGTTTCCATGGTGCATTTTGAANAATGACAGAAACCCAGCTGGAGAATAC	480						
Db	421	AACAGAGAGTGTGTTTCCATGGTGCATTTTGAANAATGACAGAAACCCAGCTGGAGAATAC	480						
Qy	481	CTACTCTTTTATTCAGAGTGAAGCTACCAATTACACAATATTGTTTACAGTCAGTATAGA	540						
Db	481	CTACTCTTTTATTCAGAGTGAAGCTACCAATTACACAATATTGTTTACAGTCAGTATAGA	540						
Qy	541	AATACCCCTGCTTTACACATTAAAGAGCTTTACTTTAGAAAAATGGAAAAACCCAGAGGCC	600						
Db	541	AATACCCCTGCTTTACACATTAAAGAGCTTTACTTTAGAAAAATGGAAAAACCCAGAGGCC	600						
Qy	601	CTGCTGTGCATATCTGAGAGCGTTCCAGAGCCGATCTGTGGAAATGGGTGCTTTGGGATTC	660						
Db	601	CTGCTGTGCATATCTGAGAGCGTTCCAGAGCCGATCTGTGGAAATGGGTGCTTTGGGATTC	660						
Qy	661	CAGGGGAAAGCTCTAAAGAGAAGTCCACGCTGTGTTTAAAAAGGAGAAAAAGTGCTT	720						
Db	661	CAGGGGAAAGCTCTAAAGAGAAGTCCACGCTGTGTTTAAAAAGGAGAAAAAGTGCTT	720						
Qy	721	CATGAATTTATTTGGACGGACATAAAGTGCTGTGCCAGAAATGAACGGGACGGGAATGC	780						
Db	721	CATGAATTTATTTGGACGGACATAAAGTGCTGTGCCAGAAATGAACGGGACGGGAATGC	780						
Qy	781	ACCAGGCTGTTTACAATAGATCTTAAATCAAACTCCTCAGACCCATTTGCCACAATTTAT	840						
Db	781	ACCAGGCTGTTTACAATAGATCTTAAATCAAACTCCTCAGACCCATTTGCCACAATTTAT	840						
Qy	841	CTTAAAGTAGGGGAACCCCTTATGATAGGTTGCAAGCTGTTTCATGTGAACCATGGATTC	900						
Db	841	CTTAAAGTAGGGGAACCCCTTATGATAGGTTGCAAGCTGTTTCATGTGAACCATGGATTC	900						
Qy	901	GGGCTCACCTCGGAAATAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960						
Db	901	GGGCTCACCTCGGAAATAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960						
Qy	961	ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGGTTTTGTATCATCAGTGCCA	1020						
Db	961	ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGGTTTTGTATCATCAGTGCCA	1020						
Qy	1021	AGAAACGACACGGGATATACACTGTTCTCTTCAAAAGCATCCGAGTCAATCAGCTTTG	1080						
Db	1021	AGAAACGACACGGGATATACACTGTTCTCTTCAAAAGCATCCGAGTCAATCAGCTTTG	1080						

1081 GTTACCATCGTAGAAAGGATTATATAATGCTACCAATCAAGTGAAGATTATGAATT 1140
1081 GTTACCATCGTAGAAAGGATTATATAATGCTACCAATTCAGTGAAGATTATGAATT 1140
1141 GACCAATATGAAGAGTTTGTCTGTGAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
1141 GACCAATATGAAGAGTTTGTCTGTGAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
1201 AGTGGACCTCTCTCGAAATCATTTCTGTGAGGTTTAAAGCCCTACCCACAAATCAGATGT 1260
1201 AGTGGACCTCTCTCGAAATCATTTCTGTGAGGTTTAAAGCCCTACCCACAAATCAGATGT 1260
1261 AGCATATCAAGTTTGTGCAATCATAGCACACGAGGAGGATATATATTTCCATGAGAA 1320
1261 AGCATATCAAGTTTGTGCAATCATAGCACACGAGGAGGATATATATTTCCATGAGAA 1320
1321 AATGATGATGCCCAATTTACCAAAATGTTTCACGCTGAATATAAGAGGAAACCTCAAGTG 1380
1321 AATGATGATGCCCAATTTACCAAAATGTTTCACGCTGAATATAAGAGGAAACCTCAAGTG 1380
1381 CTCGAGAGCATCGCAGTACAGGCGTCTGTCTCGATGGATACCCATTTACCATCT 1440
1381 CTCGAGAGCATCGCAGTACAGGCGTCTGTCTCGATGGATACCCATTTACCATCT 1440
1441 TGSACCTGAAAGAGTTTCAGACAGTCTCCAACTGCACAGAGATCAGAGGA 1500
1441 TGSACCTGAAAGAGTTTCAGACAGTCTCCAACTGCACAGAGATCAGAGGA 1500
1501 GTCTGGAATAGAAAGGCTAACGAAAGTTTGGACAGTGGGTGTCGACAGTACTCTA 1560
1501 GTCTGGAATAGAAAGGCTAACGAAAGTTTGGACAGTGGGTGTCGACAGTACTCTA 1560
1561 ACATGAGTGAAGCCATAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620
1561 ACATGAGTGAAGCCATAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620
1621 ACATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTCATCCAGACAAC 1680
1621 ACATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTCATCCAGACAAC 1680
1681 ATCTCATCTATGCAAAATGGTGTGTCTCTCTCATTTGTCGTTTAAACCCCTGCTA 1740
1681 ATCTCATCTATGCAAAATGGTGTGTCTCTCTCATTTGTCGTTTAAACCCCTGCTA 1740
1741 ATTTGTCACAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCACAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCTCAGATAATCAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860
1801 ACCGGCTCTCAGATAATCAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860
1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
1921 GGAAGTATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
1921 GGAAGTATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
1981 GCCGTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTTCAGAA 2040
1981 GCCGTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTTCAGAA 2040
2041 CTCAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGGTGC 2100
2041 CTCAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGGTGC 2100
2101 ACACGTCTAGGACCAATTTACTTGTATTTTGAATCTGTGTGATCTTCTCAAC 2160
2101 ACACGTCTAGGACCAATTTACTTGTATTTTGAATCTGTGTGATCTTCTCAAC 2160

2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
2221 AATTTCAAGTTTACCCTACTTTCCAATCATACATCCAAATTCACGATGCCCTGGTTCAAGA 2280
2221 AATTTCAAGTTTACCCTACTTTCCAATCATACATCCAAATTCACGATGCCCTGGTTCAAGA 2280
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
2341 TCTGAAGATCAAAATTTGAATATGAATAACCAAAAGGCTTGAAGAGAGGAGGACTTTGAAT 2400
2341 TCTGAAGATCAAAATTTGAATATGAATAACCAAAAGGCTTGAAGAGAGGAGGACTTTGAAT 2400
2401 GTGCTTACATTTGAAGATCTTCTTTGGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2401 GTGCTTACATTTGAAGATCTTCTTTGGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2461 CTGGAATTTAAGTCGTGTTCACAGACACTGCGCCGACGAGTCTTGTTCACCCAC 2520
2461 CTGGAATTTAAGTCGTGTTCACAGACACTGCGCCGACGAGTCTTGTTCACCCAC 2520
2521 GGGAAAGTGTGAAGATATGTACTTTGGATTGGCTCGAGATATCATCAGTATGATCCAAAC 2580
2521 GGGAAAGTGTGAAGATATGTACTTTGGATTGGCTCGAGATATCATCAGTATGATCCAAAC 2580
2581 TATGTTCTCAGGGGCAATGCCCTGCTGCTGTTAAATGGATGGCCCCCGAAGGCTGT 2640
2581 TATGTTCTCAGGGGCAATGCCCTGCTGCTGTTAAATGGATGGCCCCCGAAGGCTGT 2640
2641 GAAGGATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTAATCTGTGGAAATC 2700
2641 GAAGGATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTAATCTGTGGAAATC 2700
2701 TTCTCACTTGTGTGAATCTTACCTGGCATTCGGTGTGATGCTTAACCTTCTCAAACTG 2760
2701 TTCTCACTTGTGTGAATCTTACCTGGCATTCGGTGTGATGCTTAACCTTCTCAAACTG 2760
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATA 2820
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATA 2820
2821 ATGCAATCTGTCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880
2821 ATGCAATCTGTCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880
2881 TTTTATAGGATGTCAGCTGGCAGATGCAAGAACCGGATGATCAGATGTTGGATGGCGGT 2940
2881 TTTTATAGGATGTCAGCTGGCAGATGCAAGAACCGGATGATCAGATGTTGGATGGCGGT 2940
2941 GTTTCGGAATGTCCTCACACTACCAAAACAGGCGACTTTTCAGCAGAGATGGATTG 3000
2941 GTTTCGGAATGTCCTCACACTACCAAAACAGGCGACTTTTCAGCAGAGATGGATTG 3000
3001 GGGCTACTCTCTCCAGGCTCAGGTGCAAGTTCTAGAGGAAACAATTTAGTTTAAAG 3060
3001 GGGCTACTCTCTCCAGGCTCAGGTGCAAGTTCTAGAGGAAACAATTTAGTTTAAAG 3060
3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAACAGAGATTAATTTCAAT 3120
3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAACAGAGATTAATTTCAAT 3120
3121 CACTAAAAGAAATCTATTATCAACTGCTCTCACCAGACTTTTCTCTAGAAGCCGCT 3180
3121 CACTAAAAGAAATCTATTATCAACTGCTCTCACCAGACTTTTCTCTAGAAGCCGCT 3180
3181 GCCTTACTCTCTTTTCAAGGGACTTTTGTAAAATCAAAATCATCTCTGCACAGGCAG 3240
3181 GCCTTACTCTCTTTTCAAGGGACTTTTGTAAAATCAAAATCATCTCTGCACAGGCAG 3240
3241 GAGGAGCTGATATGAAGTATTTATTTGAGGATTTGATGTGCATCCAAAGGCTTCTCAGGCCG 3300

Db 3241 GAGGAGCTGATAATGAACCTTTATGGAGCAATTGATCTGCATCCAGGCGCTTCTCAGGCGC 3300
Qy 3301 GCTTGAGTGAATGTTGTTACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAGC 3360
Db 3301 GCTTGAGTGAATGTTGTTACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAGC 3360
Qy 3361 ATTTTGTCTAAGGAGAAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAAA 3420
Db 3361 ATTTTGTCTAAGGAGAAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAAA 3420
Qy 3421 TTTTTCAGCTATTTAGTGTATATATTTATGGTGGGAATAAAATTTCTACTACAGA 3476
Db 3421 TTTTTCAGCTATTTAGTGTATATATTTATGGTGGGAATAAAATTTCTACTACAGA 3476

RESULT 13
US-08-222-616-22
; Sequence 22, Application US/08222616
; Patent No 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935
; FILING DATE: 22-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 821P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3120 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-222-616-22

Query Match 88.1%; Score 3083.2; DB 1; Length 3120;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3097; Conservative 0; Mismatches . 23; Indels 0; Gaps 0;
Qy 58 ATCCCGGGGTTGGCGCGGACCGGCGCGCTGCTGCTGTTTCTGCAATG 117

Db 1 ATGAGAGCGTTGGCGCGGACGCGCGCCAGCTGCGCTGCTGTTGTTTTCTGCAATG 60
Qy 118 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177
Db 61 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120
Qy 178 AAGAAATATGATTCATCAGTGGGAAGTATCATCATATATCCATGGTATCAGAAATCCCG 237
Db 121 AAGAAATATGATTCATCAGTGGGAAGTATCATCATATATCCATGGTATCAGAAATCCCG 180
Qy 238 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACAAAGCTGCGCT 297
Db 181 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACAAAGCTGCGCT 240
Qy 298 GTGGAAGTGGATGTATCTGCTTCCATCACATGCAAGTGTGCTGATGCCCCAGGGAAC 357
Db 241 GTGGAAGTGGATGTATCTGCTTCCATCACATGCAAGTGTGCTGATGCCCCAGGGAAC 300
Qy 358 ATTTCTCTGCTCTGGGTCTTTAAGCACAGCTCCCTGAATGGCCAGCACATTTTGTATTA 417
Db 301 ATTTCTCTGCTCTGGGTCTTTAAGCACAGCTCCCTGAATGGCCAGCACATTTTGTATTA 360
Qy 418 CAAAACAGAGGAGTGTGTTTCCATGGTCAATTTCAAAATGACAGAAACCCAGCTGGAGAA 477
Db 361 CAAAACAGAGGAGTGTGTTTCCATGGTCAATTTGAAAATGACAGAAACCCAGCTGGAGAA 420
Qy 478 TACCTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTGTATTA 537
Db 421 TACCTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTGTATTA 480
Qy 538 AGAAATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAC 597
Db 481 AGAAATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAC 540
Qy 598 GCCCTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAGGGTGTCTTTGGCAT 657
Db 541 GCCCTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATCTGGAATGGTGTCTTTGGCAT 600
Qy 658 TCACAGGGGGAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGAGGAGGAAAAAGTG 717
Db 601 TCACAGGGGGAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGAGGAGGAAAAAGTG 660
Qy 718 CTTATGATTAATTTTGGGACGGACATAAGTGTCTGCCAGAAATGAACCTGGGCGAGGAA 777
Db 661 CTTATGATTAATTTTGGGACGGACATAAGTGTCTGCCAGAAATGAACCTGGGCGAGGAA 720
Qy 778 TGCACACAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACCAACATTTGCCACAATTA 837
Db 721 TGCACACAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACCAACATTTGCCACAATTA 780
Qy 838 TTTCTTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTTCAATGTGAACCATGGA 897
Db 781 TTTCTTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTTCAATGTGAACCATGGA 840
Qy 898 TTCGGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 957
Db 841 TTCGGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 900
Qy 958 AGTACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCATCAGTG 1017
Db 901 AGTACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCATCAGTG 960
Qy 1018 GCAAGAAACGACACCGGATACTACACTGTTCTCTTTCAAGAGTCCAGCTCAATCAGCT 1077
Db 961 GCAAGAAACGACACCGGATACTACACTGTTCTCTTTCAAGAGTCCAGCTCAATCAGCT 1020
Qy 1078 TTTGTTACCATCGTAGGAAAGGGATTTTAAATGCTTACCATTCACGTGAAGATTTATGAA 1137
Db 1021 TTTGTTACCATCGTAGGAAAGGGATTTTAAATGCTTACCATTCACGTGAAGATTTATGAA 1080
Qy 1138 ATTGACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCTTAAAGCTACCCCAACAATCAGA 1197

Db 1081 ATTGACCAATATGAAGAGTTTTGTTTTCTGTCAAGTTTAAAGCCTACCCACAATACAGA 1140
Qy 1198 TGTACGTGGACCTTCTCTCGAAAAATCATTTCTTGTGTGAGCAAAAGGGTCTTGATAACGGA 1257
Db 1141 TGTACGTGGACCTTCTCTCGAAAAATCATTTCTTGTGTGAGCAAAAGGGTCTTGATAACGGA 1200
Qy 1258 TACAGCATATCCAAAGTTTTGCAATATGAAGCAGCCAGGAGAGATATATATCCATGCA 1317
Db 1201 TACAGCATATCCAAAGTTTTGCAATATGAAGCAGCCAGGAGAGATATATATCCATGCA 1260
Qy 1318 GAAATGATGATGCCAATTTACCAAAATGTTACAGCTGATATATGAAGAAACCTCA 1377
Db 1261 GAAATGATGATGCCAATTTACCAAAATGTTACAGCTGATATATGAAGAAACCTCA 1320
Qy 1378 GTGCTCCAGAAAGCATCGGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATTACCA 1437
Db 1321 GTCTCCAGAAAGCTTCGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATTACCA 1380
Qy 1438 TCTTGGACCTTGAAGAGTGTTCAGCAAGTCTCCCAAGTGCACAGAGAGATCACAGAA 1497
Db 1381 TCTTGGACCTTGAAGAGTGTTCAGCAAGTCTCCCAAGTGCACAGAGAGATCACAGAA 1440
Qy 1498 GGAGTCTGGAATGAAGAGCTAACAGAAAGTGTTCGACAGTGGGTGTCAGCAGTACT 1557
Db 1441 GGAGTCTGGAATGAAGAGCTAACAGAAAGTGTTCGACAGTGGGTGTCAGCAGTACT 1500
Qy 1558 CTAACATGATGTAAGGCATATAAAGGGTCTCTGCTCAAGTCTGTGCATACAAATCCCTT 1617
Db 1501 CTAACATGATGTAAGGCATATAAAGGGTCTCTGCTCAAGTCTGTGCATACAAATCCCTT 1560
Qy 1618 GGCACATCTTGTGAGAGATCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1677
Db 1561 GGCACATCTTGTGAGAGATCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1620
Qy 1678 AACATCTCATCTATGCAACAATTTGGTGTGTCTCTCTCTCTCATTTGCTGTTTTAACCCCTG 1737
Db 1621 AACATCTCATCTATGCAACAATTTGGTGTGTCTCTCTCTCTCATTTGCTGTTTTAACCCCTG 1680
Qy 1738 CTAATTTGTCACAGATACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATCGTACAG 1797
Db 1681 CTAATTTGTCACAGATACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATCGTACAG 1740
Qy 1798 GTGACCGCTCTCTCAGATATAGTACTCTCTAGCTTGATTTACAGAGATATGAATATGAT 1857
Db 1741 GTGACCGGATCTCTCAGATATAGTACTCTCTAGCTTGATTTACAGAGATATGAATATGAT 1800
Qy 1858 CTCAAATGGAGTTTCCAGAGAAATTTAGTTTGGGAGGTACTAGGATCAGGTGCT 1917
Db 1801 GTCAAATGGAGTTTCCAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCT 1860
Qy 1918 TTTGGAAAAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977
Db 1861 TTTGGAAAAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920
Qy 1978 GTTGGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
Db 1921 GTTACCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
Qy 2038 GAACTCAAGATGATGCCAGCTGGGAGCCAGCAGAAATTTGTGAACCTGCTGGGGCG 2097
Db 1981 GAACTCAAGATGATGCCAGCTGGGAGCCAGCAGAAATTTGTGAACCTGCTGGGGCG 2040
Qy 2098 TGCACACTCTCAGGACCAATTTACTTGATTTTGAATAGTGTGCTATGATGATCTCTC 2157
Db 2041 TGCACACTCTCAGGACCAATTTACTTGATTTTGAATAGTGTGCTATGATGATCTCTC 2100
Qy 2158 AACTATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGACAGAGATTTTCAAGSAA 2217
Db 2101 AACTATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGACAGAGATTTTCAAGSAA 2160
Qy 2218 CACAATTTTCAGTTTTTACCCCACTTTCCCAATCATCTCAATTTCCAGCATGCTGGTTC 2277
Db 2161 CACAATTTTCAGTTTTTACCCCACTTTCCCAATCATCTCAATTTCCAGCATGCTGGTTC 2220

Qy 2278 AGAAGAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGAAATTCATTT 2337
Db 2221 AGAAGAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGAAATTCATTT 2280
Qy 2338 CACTCTGAAGATGAAATTTGAATATGAACACCAAAAAGGCTGGAAGAGAGGAGGACTTG 2397
Db 2281 CACTCTGAAGATGAAATTTGAATATGAACACCAAAAAGGCTGGAAGAGAGGAGGACTTG 2340
Qy 2398 AATGCTTACATTTGAAGATCTTCTTGTGCTTGGATATCAAGTTGCCAAAGAAATGGAA 2457
Db 2341 AATGCTTACATTTGAAGATCTTCTTGTGCTTGGATATCAAGTTGCCAAAGAAATGGAA 2400
Qy 2458 TTTCTGGAAATTTAAGTCGTGTTTACAGAGACCTGGCGCCAGGAACTGCTTGTGCACC 2517
Db 2401 TTTCTGGAAATTTAAGTCGTGTTTACAGAGACCTGGCGCCAGGAACTGCTTGTGCACC 2460
Qy 2518 CACGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTTCGAGATATCATGATGATTC 2577
Db 2461 CACGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTTCGAGATATCATGATGATTC 2520
Qy 2578 AACTATGTTGTGACGGGCAATGCCGTCTGCTGTAAATGATGGCCCCGAAAGCCCTG 2637
Db 2521 AACTATGTTGTGACGGGCAATGCCGTCTGCTGTAAATGATGGCCCCGAAAGCCCTG 2580
Qy 2638 TTTGAAGCATCTACACCATTAAGAGTGTCTGCTGATATGGAATATTACTGTGGGAA 2697
Db 2581 TTTGAAGCATCTACACCATTAAGAGTGTCTGCTGATATGGAATATTACTGTGGGAA 2640
Qy 2698 ATCTTCTCACATTTGGTGAATCTTACCTTGGCTTACCTTGGCTTACCTTGGCTTACCTTACAAA 2757
Db 2641 ATCTTCTCACATTTGGTGAATCTTACCTTGGCTTACCTTGGCTTACCTTACAAA 2700
Qy 2758 CTGATTTCAAAATGGAATTTAAATGATGATCAGCCATTTATGCTTACAGAGAAATATACAT 2817
Db 2701 CTGATTTCAAAATGGAATTTAAATGATGATCAGCCATTTTATGCTTACAGAGAAATATACAT 2760
Qy 2818 ATAATGCAATCTCTGCTGGCTTTTACCTCAAGGAAAGGCCATCTCTCCCTAATTTGACT 2877
Db 2761 ATAATGCAATCTCTGCTGGCTTTTACCTCAAGGAAAGGCCATCTCTCCCTAATTTGACT 2820
Qy 2878 TCGTTTTTAGGATGTGCTGCTGAGATGCAAGAAAGCGATGTATCAGAACTGTGATGCG 2937
Db 2821 TCGTTTTTAGGATGTGCTGCTGAGATGCAAGAAAGCGATGTATCAGAACTGTGATGCG 2880
Qy 2938 CGTGTTCGGAATGTCTTACACCTTACCAAAACAGGCGACCTTTTACAGAGAGATGAT 2997
Db 2881 CGTGTTCGGAATGTCTTACACCTTACCAAAACAGGCGACCTTTTACAGAGAGATGAT 2940
Qy 2998 TTTGGGCTACTCTCTCCGAGCTCAGTCTGAAGATTCTGAGAGAAATTTAGTTT 3057
Db 2941 TTTGGGCTACTCTCTCCGAGCTCAGTCTGAAGATTCTGAGAGAAATTTAGTTT 3000
Qy 3058 AGGACTTCACTCCCTCCACCTATCCCTAACAGCTGTAGATTACCAAAACAGATTAATTT 3117
Db 3001 AGGACTTCACTCCCTCCACCTATCCCTAACAGCTGTAGATTACCAAAACAGATTAATTT 3060
Qy 3118 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACAGAGCTTTTCTCTAGAGCGG 3177
Db 3061 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACAGAGCTTTTCTCTAGAGCGG 3120

RESULT 14

US-08-446-648-22
; Sequence 22, Application US/08446648
; Patent No. 6331302
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping

QY 1618 GGCACATCTTGTGAGACGATCCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1677
DB 1561 GGCACATCTTGTGAGACGATCCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1620
QY 1678 AACATCTCATCTATGATCAACAATTTGGTGTGTGCTCTCTTCATGTCTGTTTAAACCGTG 1737
DB 1621 AACATCTCATCTATGATCAACAATTTGGTGTGTGCTCTCTTCATGTCTGTTTAAACCGTG 1680
QY 1738 CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATGATG 1797
DB 1681 CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATGATG 1740
QY 1798 GTACCGGCTTCTCAGATATGATGATCTTACGTTGATTTTCAAGAGATATGATATGAT 1857
DB 1741 GTACCGGATCTCAGATATGATGATCTTACGTTGATTTTCAAGAGATATGATATGAT 1800
QY 1858 CTCAAATGGGAGTTTCCACAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCT 1917
DB 1801 GTCAAATGGGAGTTTCCACAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCT 1860
QY 1918 TTTGGAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977
DB 1861 TTTGGAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920
QY 1978 GTTCCGCTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
DB 1921 GTTACCGCTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
QY 2038 GAACTCAAGATGATGACCCAGCTGGGAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2097
DB 1981 GAACTCAAGATGATGACCCAGCTGGGAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2040
QY 2098 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATACGTTGCTATGATGATCTCTC 2157
DB 2041 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATACGTTGCTATGATGATCTCTC 2100
QY 2158 AACTATCTAAGAAAGTAAAGAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2217
DB 2101 AACTATCTAAGAAAGTAAAGAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2160
QY 2218 CACAATTTCAAGTTTACCCACCTTTCCATCATCATCAATTTCCAGCATGCTGTTTCA 2277
DB 2161 CACAATTTCAAGTTTACCCACCTTTCCATCATCATCAATTTCCAGCATGCTGTTTCA 2220
QY 2278 AGAAGATTCAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTT 2337
DB 2221 AGAAGATTCAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTT 2280
QY 2338 CACTCTGAAGATGAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGAGGACTTG 2397
DB 2281 CACTCTGAAGATGAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGAGGACTTG 2340
QY 2398 AATGTGCTTACATTTGAAGATCTTCTTGTGCTATCAAGTTCGCAAGGAATGGAA 2457
DB 2341 AATGTGCTTACATTTGAAGATCTTCTTGTGCTATCAAGTTCGCAAGGAATGGAA 2400
QY 2458 TTTCTGGAATTTAAGTCGTGTGTACAGAGACCTGCCGCCAGGACGCTGTTGCACC 2517
DB 2401 TTTCTGGAATTTAAGTCGTGTGTACAGAGACCTGCCGCCAGGACGCTGTTGCACC 2460
QY 2518 CACGGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTC 2577
DB 2461 CACGGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTC 2520
QY 2578 AACTATGTTGTACGGGGCAATGCCCGCTGCTGCTGTTAAATGATGAGGCCCCCAAGCCGTG 2637
DB 2521 AACTATGTTGTACGGGGCAATGCCCGCTGCTGCTGTTAAATGATGAGGCCCCCAAGCCGTG 2580
QY 2638 TTTGAAGGATCTACACCAATTAAGAGTGTGCTGCTATGATATGATATGATGATGATG 2697
DB 2581 TTTGAAGGATCTACACCAATTAAGAGTGTGCTGCTATGATATGATATGATGATGATG 2640
QY 2698 ATCTTCTACTTGGTGTGAATCTTACCCTGGCATTCCGCTTGCATCTCACTTCTACAAA 2757

DB 2641 ATCTTCTACATTTGGTGTGAATCCCTTACCTTGGCATTCGCGTTGATGCTTACTTCTACAAA 2700
QY 2758 CTGATTTCAAAATGATTTAAATGATGATCAGCCATTTTATGCTACAGAAAGAAATATACATT 2817
DB 2701 CTGATTTCAAAATGATTTAAATGATGATCAGCCATTTTATGCTACAGAAAGAAATATACATT 2760
QY 2818 ATAATGAATCCTCTCGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAAATTTGACT 2877
DB 2761 ATAATGAATCCTCTCGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAAATTTGACT 2820
QY 2878 TCGTTTTTAGGATGTCTGAGTGGCAGATGCAGAAAGAGCGATGTATCAGAAATGTGGATGCG 2937
DB 2821 TCGTTTTTAGGATGTCTGAGTGGCAGATGCAGAAAGAGCGATGTATCAGAAATGTGGATGCG 2880
QY 2938 CGTGTTCGGAATGTCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2997
DB 2881 CGTGTTCGGAATGTCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2940
QY 2998 TTGGGGTACTCTCTCGCAGGCTCAGTCCGAAAGTTCGTAGAGAAACAAATTTAGTTTAA 3057
DB 2941 TTGGGGTACTCTCTCGCAGGCTCAGTCCGAAAGTTCGTAGAGAAACAAATTTAGTTTAA 3000
QY 3058 AGGACTTCTATCCCTCCACCTATCCCTACAGGCTGTAGATTACCAAAAACAGATTAATTT 3117
DB 3001 AGGACTTCTATCCCTCCACCTATCCCTACAGGCTGTAGATTACCAAAAACAGATTAATTT 3060
QY 3118 CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGAGCTTTTCTCTAGAGCGC 3177
DB 3061 CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGAGCTTTTCTCTAGAGCGC 3120

RESULT 15

PCT-US95-04228-22
; Sequence 22, Application PC/TUS9504228
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04228
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/22616
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 821P3PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 3120 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

PCT-US95-04228-22

Query Match 88.1%; Score 3083.2; DB 5; Length 3120;

Best Local Similarity 99.3%; Pred. No. 0;
Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY	58	ATGCCGCGGTTGGCGCGGACCGGCGCGCTGCTGCTGTTTTTCTGCAATG	117
Db	1	ATGAGAGCGTTGGCGCGGACGCGCGCTGCTGCTGTTTTTCTGCAATG	60
QY	118	ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCATCAT	177
Db	61	ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCATCAT	120
QY	178	AGAACAAATGATTCATCAGTGGGAAAGTCATCATCATATCCCATGTATCAGAAATCCCG	237
Db	121	AGAACAAATGATTCATCAGTGGGAAAGTCATCATCATATCCCATGTATCAGAAATCCCG	180
QY	238	GAAGACCTCGGCTGCTGCTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCT	297
Db	181	GAAGACCTCGGCTGCTGCTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCT	240
QY	298	GTGGAAGTGGATGTATCTGCTCCATCAGACTGCAAGTGTGTCGATGCCCGCAGGAAC	357
Db	241	GTGGAAGTGGATGTATCTGCTCCATCAGACTGCAAGTGTGTCGATGCCCGCAGGAAC	300
QY	358	ATTTCCTGCTCTGGCTTTTAAAGACAGCTCCCTGAATGCCAGCCACATTTTGATTTA	417
Db	301	ATTTCCTGCTCTGGCTTTTAAAGACAGCTCCCTGAATGCCAGCCACATTTTGATTTA	360
QY	418	CAAAACAGAGGAGTGTGCTCCATGTGCTATTTGAAATGACAGAAACCCCAAGCTGGAGAA	477
Db	361	CAAAACAGAGGAGTGTGCTCCATGTGCTATTTGAAATGACAGAAACCCCAAGCTGGAGAA	420
QY	478	TACCTACTTTTATTACAGAGTGAAGCTACCAATTTACACAATATTTTACAGTACTATA	537
Db	421	TACCTACTTTTATTACAGAGTGAAGCTACCAATTTACACAATATTTTACAGTACTATA	480
QY	538	AGAAATACCTGCTTTACACATTAAGACAGCTTACTTTAGAAAAATGGAACCCAGGAC	597
Db	481	AGAAATACCTGCTTTACACATTAAGACAGCTTACTTTAGAAAAATGGAACCCAGGAC	540
QY	598	GCCCTGCTGCTATCTGAGAGCTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	657
Db	541	GCCCTGCTGCTATCTGAGAGCTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	600
QY	658	TCACAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG	717
Db	601	TCACAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG	660
QY	718	CTTCATCAATTTTGGGAGCGATAGTGTGCTGCCAGAAATGAACCTGGGCGAGGAA	777
Db	661	CTTCATCAATTTTGGGAGCGATAGTGTGCTGCCAGAAATGAACCTGGGCGAGGAA	720
QY	778	TGCACACGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCCACATTTGCCACAATTA	837
Db	721	TGCACACGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCCACATTTGCCACAATTA	780
QY	838	TTTCTTTAAAGTAGGGAACCTTTATGGATAAGTGTGCAAAAGCTGTTATGTGAACCATGGA	897
Db	781	TTTCTTTAAAGTAGGGAACCTTTATGGATAAGTGTGCAAAAGCTGTTATGTGAACCATGGA	840
QY	898	TTCCGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCGCACTACTTTGAGATG	957
Db	841	TTCCGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCGCACTACTTTGAGATG	900
QY	958	AGTACCTATTCAACAACACAACATATGATACGGATCTGTTGCTTTTGTATCATCATG	1017
Db			

Db	901	AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCATG	960
QY	1018	GCAAGAAACGACACCGGATACTACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCT	1077
Db	961	GCAGAAACGACACCGGATACTACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCT	1020
QY	1078	TTGGTTACCATCGTAGGAAAGGATTTATAAAATGCTACCAATTCAGTGAAGATATGAA	1137
Db	1021	TTGGTTACCATCGTAGGAAAGGATTTATAAAATGCTACCAATTCAGTGAAGATATGAA	1080
QY	1138	ATTGACCAATATCAAGAGTTTTTCTGCTCAGGTTTAAAGCTACCCACAATCAGA	1197
Db	1081	ATTGACCAATATCAAGAGTTTTTCTGCTCAGGTTTAAAGCTACCCACAATCAGA	1140
QY	1198	TGTACGTGGACCTCTCTCGAAAATCATTTCCCTGTGAGCAAAAGGCTCTTTGTAACGGA	1257
Db	1141	TGTACGTGGACCTCTCTCGAAAATCATTTCCCTGTGAGCAAAAGGCTCTTTGTAACGGA	1200
QY	1258	TACAGCATATCCAAGTTTTGCAATCATPAAGCACAGCCAGAGAGATATATATCCATGCA	1317
Db	1201	TACAGCATATCCAAGTTTTGCAATCATPAAGCACAGCCAGAGAGATATATATCCATGCA	1260
QY	1318	GAAATGATGATGCCCAATTTACCAAAATGTTCCAGCTGATATAAGAAAGAAACCTCAA	1377
Db	1261	GAAATGATGATGCCCAATTTACCAAAATGTTCCAGCTGATATAAGAAAGAAACCTCAA	1320
QY	1378	GTCTGCGACAAGCATCGCAAGTCAGCGCTCCCTGTTTCTCGGATGGATACCCATTACCA	1437
Db	1321	GTCTGCGACAAGCATCGCAAGTCAGCGCTCCCTGTTTCTCGGATGGATACCCATTACCA	1380
QY	1438	TCTTGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAAGTGCACAGAGAGATCACAGAA	1497
Db	1381	TCTTGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAAGTGCACAGAGAGATCACAGAA	1440
QY	1498	GGAGTCTGGATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGGTGTCAGCAGTACT	1557
Db	1441	GGAGTCTGGATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGGTGTCAGCAGTACT	1500
QY	1558	CTAAACATGATGAAGCCATAAAAGGTTCTCTGCTCAAGTCTGTGCATACAAATCCCTT	1617
Db	1501	CTAAACATGATGAAGCCATAAAAGGTTCTCTGCTCAAGTCTGTGCATACAAATCCCTT	1560
QY	1618	GGCACAATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC	1677
Db	1561	GGCACAATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC	1620
QY	1678	AACATCTATCTATGACAAATTTGGTGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1737
Db	1621	AACATCTATCTATGACAAATTTGGTGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1680
QY	1738	CTAATTTGTCACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
Db	1681	CTAATTTGTCACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
QY	1798	GTGACCGGCTCTCAGATATAGTACTTCTACGTTGATTTTTCAGAGAAATATGAATATGAT	1857
Db	1741	GTGACCGGCTCTCAGATATAGTACTTCTACGTTGATTTTTCAGAGAAATATGAATATGAT	1800
QY	1858	CTCAATTTGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGGTTACTAGGATCAGGTGCT	1917
Db	1801	CTCAATTTGGGAGTTTCCAAGAGAAATTTAGAGTTTGGGAAGGTTACTAGGATCAGGTGCT	1860
QY	1918	TTTGGAAAGTGTAGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG	1977
Db	1861	TTTGGAAAGTGTAGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG	1920
QY	1978	GTTCCCTGCTAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	2037
Db	1921	GTTCCCTGCTAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	1980
QY	2038	GAATCTAAGATGATGACCCAGCTGGGAAGCCAGCAAGATATTTGTGAACCTGCTGGGGGG	2097
Db	1981	GAATCTAAGATGATGACCCAGCTGGGAAGCCAGCAAGATATTTGTGAACCTGCTGGGGGG	2040

Search completed: May 25, 2003, 12:20:06
Job time : 150.504 secs

```
QY 2098 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATACTGTTGCTATGGTGATCTTC 2157
Db 2041 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATACTGTTGCTATGGTGATCTTC 2100
QY 2158 AACTATCTAAGAAGTAAAGAGAAATAATTCACAGGACTTGGACAGAGATTTTCAAGGAA 2217
Db 2101 AACTATCTAAGAAGTAAAGAGAAATAATTCACAGGACTTGGACAGAGATTTTCAAGGAA 2160
QY 2218 CACAATTTCACTTTTACCCACATTTCCAAATCACATCCAAATTCAGCATGCCCTGGTTCA 2277
Db 2161 CACAATTTCACTTTTACCCACATTTCCAAATCACATCCAAATTCAGCATGCCCTGGTTCA 2220
QY 2278 AGAGAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2337
Db 2221 AGAGAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2280
QY 2338 CACTCTGAAGATGAATTAATATGAATAACCAAAAGGCTGGAAGAGAGAGGACTTG 2397
Db 2281 CACTCTGAAGATGAATTAATATGAATAACCAAAAGGCTGGAAGAGAGAGGACTTG 2340
QY 2398 AATGCTGTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTCGCAAGGAATGGAA 2457
Db 2341 AATGCTGTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTCGCAAGGAATGGAA 2400
QY 2458 TTTCTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCGGCAGCAAGCTGCTGTCAAC 2517
Db 2401 TTTCTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCGGCAGCAAGCTGCTGTCAAC 2460
QY 2518 CACGGGAAAGTGTGAAGATATGACTTTTGGATTGGCTCGAGATATCATGAGTGATTC 2577
Db 2461 CACGGGAAAGTGTGAAGATATGACTTTTGGATTGGCTCGAGATATCATGAGTGATTC 2520
QY 2578 AACTATGTTGTCAGGGGGAATGCCGCTGCTGTAAATGGATGGCCCCCGAAAGCCTG 2637
Db 2521 AACTATGTTGTCAGGGGGAATGCCGCTGCTGTAAATGGATGGCCCCCGAAAGCCTG 2580
QY 2638 TTTGAAGGCATCTACACCATTAAGAGTGATCTCTGGTCATATGGAATATTACTGTGGAA 2697
Db 2581 TTTGAAGGCATCTACACCATTAAGAGTGATCTCTGGTCATATGGAATATTACTGTGGAA 2640
QY 2698 ATCTTCTCACTTGGTGTGAATCCCTTACCCTGGCATTCGGTTGATGTACTTAACCTTCAAAA 2757
Db 2641 ATCTTCTCACTTGGTGTGAATCCCTTACCCTGGCATTCGGTTGATGTACTTAACCTTCAAAA 2700
QY 2758 CTGATTCAAATGGATTTAAATGGATCAGCCATTTATGCTACAGAAGAAATATACATT 2817
Db 2701 CTGATTCAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATT 2760
QY 2818 ATATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACT 2877
Db 2761 ATATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACT 2820
QY 2878 TCGTTTTTAGGATGTACAGTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGATGCG 2937
Db 2821 TCGTTTTTAGGATGTACAGTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGATGCG 2880
QY 2938 CGTGTTCGGAATGCTCTACACCTTACCAAAACAGGCGACCTTTCAGCAGAGAGATGGAT 2997
Db 2881 CGTGTTCGGAATGCTCTACACCTTACCAAAACAGGCGACCTTTCAGCAGAGAGATGGAT 2940
QY 2998 TTGGGGCTACTCTCTCGGAGGCTCAGGTCGAGATTTCGTAGAGGAACAATTTAGTTTTTA 3057
Db 2941 TTGGGGCTACTCTCTCGGAGGCTCAGGTCGAGATTTCGTAGAGGAACAATTTAGTTTTTA 3000
QY 3058 AGGACTTCAPCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGATTAATTT 3117
Db 3001 AGGACTTCAPCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGTTAATTT 3060
QY 3118 CATCACTAAAGAAATCTATTATCAACTGCTGCTCACCAGACTTTTCTCTAGAACCG 3177
Db 3061 CATCACTAAAGAAATCTATTATCAACTGCTGCTCACCAGACTTTTCTCTAGAGCG 3120
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 11:51:23 ; Search time 403.768 Seconds
(without alignments)
11449.495 Million cell updates/sec

Title: US-09-919-408-3

Perfect score: 3501

Sequence: 1 CGAGGCGCATCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC

Gapop 10.0 , capext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3501	100.0	3501	10	US-09-919-408-3
2	3501	100.0	3501	10	US-09-872-136-3
3	3083.2	88.1	3120	10	US-09-982-610-22
4	2227.2	63.6	3453	10	US-09-919-408-1
5	2227.2	63.6	3453	10	US-09-872-136-1
6	633	18.1	645	9	US-09-796-692-7538
7	633	18.1	645	9	US-10-040-862-7538
8	608.4	17.4	610	9	US-09-796-692-8085
9	608.4	17.4	610	9	US-10-040-862-8085
10	606.4	17.3	608	9	US-09-796-692-8132
11	606.4	17.3	608	9	US-10-040-862-8132
12	605	17.3	605	9	US-09-796-692-8927
13	605	17.3	605	9	US-10-040-862-8927
14	598.4	17.1	601	9	US-09-796-692-8273
15	598.4	17.1	601	9	US-10-040-862-8273
16	597	17.1	597	9	US-09-796-692-7976
17	597	17.1	597	9	US-10-040-862-7976
18	595.4	17.0	597	9	US-09-796-692-9010
19	595.4	17.0	597	9	US-10-040-862-9010

20	593.2	16.9	598	9	US-09-796-692-8566	Sequence 8566, Ap
21	593.2	16.9	598	9	US-10-040-862-8566	Sequence 8566, Ap
22	590.8	16.9	594	9	US-09-796-692-8135	Sequence 8135, Ap
23	590.8	16.9	594	9	US-10-040-862-8135	Sequence 8135, Ap
24	590.4	16.9	605	9	US-09-796-692-8047	Sequence 8047, Ap
25	590.4	16.9	605	9	US-10-040-862-8047	Sequence 8047, Ap
26	583.8	16.7	589	9	US-09-796-692-8630	Sequence 8630, Ap
27	583.8	16.7	589	9	US-10-040-862-8630	Sequence 8630, Ap
28	562.4	16.1	564	9	US-09-796-692-9025	Sequence 9025, Ap
29	562.4	16.1	564	9	US-10-040-862-9025	Sequence 9025, Ap
30	554.4	15.8	556	9	US-09-796-692-7883	Sequence 7883, Ap
31	554.4	15.8	556	9	US-10-040-862-7883	Sequence 7883, Ap
32	541	15.5	550	9	US-09-796-692-8494	Sequence 8494, Ap
33	541	15.5	550	9	US-10-040-862-8494	Sequence 8494, Ap
34	539	15.4	539	9	US-09-796-692-8255	Sequence 8255, Ap
35	539	15.4	539	9	US-10-040-862-8255	Sequence 8255, Ap
36	510	14.6	510	9	US-09-796-692-8268	Sequence 8268, Ap
37	510	14.6	510	9	US-10-040-862-8268	Sequence 8268, Ap
38	502.8	14.4	508	9	US-09-796-692-8032	Sequence 8032, Ap
39	502.8	14.4	508	9	US-10-040-862-8032	Sequence 8032, Ap
40	491.4	14.0	497	9	US-09-796-692-8408	Sequence 8408, Ap
41	491.4	14.0	497	9	US-10-040-862-8408	Sequence 8408, Ap
42	478.4	13.7	606	9	US-09-796-692-8114	Sequence 8114, Ap
43	478.4	13.7	606	9	US-10-040-862-8114	Sequence 8114, Ap
44	446.4	12.8	581	9	US-09-796-692-9031	Sequence 9031, Ap
45	446.4	12.8	581	9	US-10-040-862-9031	Sequence 9031, Ap

ALIGNMENTS

RESULT 1

US-09-919-408-3

; Sequence 3, Application US/09919408

; Patent No. US20020072077A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ImClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/919,408

; FILING DATE: 31-Jul-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US 07/906,397

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: US PCT/US92/05401

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: TW 81102961

; FILING DATE: 15-APR-1992

; APPLICATION NUMBER: US PCT/US92/02750

; FILING DATE: 02-APR-1992

; APPLICATION NUMBER: US 07/813,593

; FILING DATE: 24-DEC-1991

; APPLICATION NUMBER: US 07/793,065

; FILING DATE: 15-NOV-1991

; APPLICATION NUMBER: US 07/728,913

; FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Peit, Irving N.
REGISTRATION NUMBER: 28 601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58...3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139...3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58...138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

us-09-919-408-3

		Query Match	Best Local Similarity	100.0%	Score 3501;	DB 10;	Length 3501;		
		Mismatches	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
QY	1	CGAGCGGCATCCGAGGGCTGGCGGCCCTCGGGGACCCCGGCTCCGGAGGCCATG	60						
DB	1	CGAGCGGCATCCGAGGGCTGGCGGCCCTCGGGGACCCCGGCTCCGGAGGCCATG	60						
QY	61	CGCGGTTGGCGGCGAGCGGGACCGGCGCTGCTGTTGTTTCTCAATGATA	120						
DB	61	CGCGGTTGGCGGCGAGCGGGACCGGCGCTGCTGTTGTTTCTCAATGATA	120						
QY	121	TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTAAATCAATCAAG	180						
DB	121	TTTGGGACTATTACAAATCAAGATCGCTGTGATCAAGTGTGTTTAAATCAATCAAG	180						
QY	181	AACAATGATTCATCAGTGGGGAAGTCATCATATATCCCATGATCAGAAATCCCGGAA	240						
DB	181	AACAATGATTCATCAGTGGGGAAGTCATCATATATCCCATGATCAGAAATCCCGGAA	240						
QY	241	GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGCACAGTGTACCAAGCTGCGCTGTG	300						
DB	241	GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGCACAGTGTACCAAGCTGCGCTGTG	300						
QY	301	GAAAGTGGATGATCTGCTTCCATCAGCTGCAAGTGTGCTGATGCCCGCAGGGAACATT	360						
DB	301	GAAAGTGGATGATCTGCTTCCATCAGCTGCAAGTGTGCTGATGCCCGCAGGGAACATT	360						
QY	361	TCCTGTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCGAGCAATTTGATTTACAA	420						
DB	361	TCCTGTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCGAGCAATTTGATTTACAA	420						
QY	421	ACACAGGAGTGTGTTCCATGCTCATTTTGAATGACAGAAACCCCAAGCTGGAGATAC	480						
DB	421	ACACAGGAGTGTGTTCCATGCTCATTTTGAATGACAGAAACCCCAAGCTGGAGATAC	480						
QY	481	CTACTTTTATTACAGAGTGAAGCTACCAATACACAAATATTGTTTACAGTGAATAAAG	540						
DB	481	CTACTTTTATTACAGAGTGAAGCTACCAATACACAAATATTGTTTACAGTGAATAAAG	540						
QY	541	AATACCCCTGTTTACACATTAAGAACCTTACTTTAGAAAAATGAAAAACCCAGGAGCC	600						
DB									

541	ATATCCCTGCTTTACACATTAAAGACACCTTACTTTAGAAAAATGAAAAACCCAGGAGCC	600
601	CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGATCGTGAATGGTGTGTTCCGATTC	660
601	CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGATCGTGAATGGTGTGTTCCGATTC	660
661	CAGGGGAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT	720
661	CAGGGGAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT	720
721	CATGAATTTTGGGACGACATAGTGTCTGTCAGAAATGAATGGGAGGAAATGC	780
721	CATGAATTTTGGGACGACATAGTGTCTGTCAGAAATGAATGGGAGGAAATGC	780
781	ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCCAGACCATGTCACAAATATT	840
781	ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCCAGACCATGTCACAAATATT	840
841	CTTAAGTAGGGGACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGATTC	900
841	CTTAAGTAGGGGACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGATTC	900
901	GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGAGGGCAACTACTTTGAGATGAGT	960
901	GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGAGGGCAACTACTTTGAGATGAGT	960
961	ACCTATTCAACAAACAGAACTATGATACGGATCTCTGTTGCTTTTGTATCATCAGTGGCA	1020
961	ACCTATTCAACAAACAGAACTATGATACGGATCTCTGTTGCTTTTGTATCATCAGTGGCA	1020
1021	AGAAAGCACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
1021	AGAAAGCACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
1081	GTTACCATCGTAGAAAGGATTATTAATGCTACCATTCAAGTGAAGATTTGAAAT	1140
1081	GTTACCATCGTAGAAAGGATTATTAATGCTACCATTCAAGTGAAGATTTGAAAT	1140
1141	GACCAATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT	1200
1141	GACCAATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT	1200
1201	ACGTGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGTAACGGATAC	1260
1201	ACGTGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGTAACGGATAC	1260
1261	AGCATATCCAAGTTTGCATATAGCACAGCCAGGAGAAATATATATCCATGCGAGAA	1320
1261	AGCATATCCAAGTTTGCATATAGCACAGCCAGGAGAAATATATATCCATGCGAGAA	1320
1321	AATGATGATGCCCAATTTACCAAAATGTTACGCTGAATATAAGAAGAAACCTCAAGTG	1380
1321	AATGATGATGCCCAATTTACCAAAATGTTACGCTGAATATAAGAAGAAACCTCAAGTG	1380
1381	CTCGAGAGCATCGGCAAGTCAAGCTGCTGTTTTCGGATGGATACCATTTACCATCT	1440
1381	CTCGAGAGCATCGGCAAGTCAAGCTGCTGTTTTCGGATGGATACCATTTACCATCT	1440
1441	TGGACCTGGAAGAGTGTTCAGACAACTGCCCAAGTCTCCCAAGTGCACAGAGAGATCACAGAGGA	1500
1441	TGGACCTGGAAGAGTGTTCAGACAACTGCCCAAGTCTCCCAAGTGCACAGAGAGATCACAGAGGA	1500
1501	GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
1501	GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
1561	AACATGAGTGAAGCCATAAAGGGTCTCTGCTGAAGTGTGTCATCAATTCCTTGGC	1620
1561	AACATGAGTGAAGCCATAAAGGGTCTCTGCTGAAGTGTGTCATCAATTCCTTGGC	1620
1621	ACATCTGTGAGAGGATCCCTTTTAACTCTCCAGGCCCCCTCCCTTTTCATCCAGACAAAC	1680
1621	ACATCTGTGAGAGGATCCCTTTTAACTCTCCAGGCCCCCTCCCTTTTCATCCAGACAAAC	1680

1681 ATCTCATCTATCAACAATTGGTGTGTTGTCCTCTCTATTGTCGTTTTAAACCCCTGCTA 1740
1681 ATCTCATCTATCAACAATTGGTGTGTTGTCCTCTCTATTGTCGTTTTAAACCCCTGCTA 1740
1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTACAGAAATATGAATATGATCTC 1860
1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTACAGAAATATGAATATGATCTC 1860
1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGATCAGGTGCTTTT 1920
1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGATCAGGTGCTTTT 1920
1921 GGAAGTGTATGAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 GGAAGTGTATGAACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 GCGGTCAAAATGCTGAAGAAAGAGAGAGTCTGAAGAGAGGAGTCTCATGTGAGAA 2040
1981 GCGGTCAAAATGCTGAAGAAAGAGAGAGTCTGAAGAGAGGAGTCTCATGTGAGAA 2040
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCTGC 2100
2101 ACACCTGTCAGGACCAATTTACTTGTATTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160
2101 ACACCTGTCAGGACCAATTTACTTGTATTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160
2161 TATCTAAGAGTAAAGAGAGAAATTTACAGAGCTTGGACAGAGATTTTCAAGAGAAC 2220
2161 TATCTAAGAGTAAAGAGAGAAATTTACAGAGCTTGGACAGAGATTTTCAAGAGAAC 2220
2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCTCAAAATTTCCAGCATGCTGTTCAAGA 2280
2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCTCAAAATTTCCAGCATGCTGTTCAAGA 2280
2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCAC 2340
2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCAC 2340
2341 TCTGAAGTAAATTTGAATATGAAGAAAGCAATTTGAGTATGAGTATGAGTATGAGT 2400
2341 TCTGAAGTAAATTTGAATATGAAGAAAGCAATTTGAGTATGAGTATGAGTATGAGT 2400
2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTCGCAAGGAATGGAATTT 2460
2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTCGCAAGGAATGGAATTT 2460
2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCCAGGAACCTGTTGTCAACCCAC 2520
2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCCAGGAACCTGTTGTCAACCCAC 2520
2521 GGAAGTGGTGAAGATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 2580
2521 GGAAGTGGTGAAGATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 2580
2581 TATGTTGTCAGGGCAATCCCGCTGCTGCTGTTAAATGATGATGATGATGATGATGATG 2640
2581 TATGTTGTCAGGGCAATCCCGCTGCTGCTGTTAAATGATGATGATGATGATGATGATG 2640
2641 GAAGGATCTACACCATTAAGAGTATGTTGTTGATATGGAATATTTACTGTGGGAATC 2700
2641 GAAGGATCTACACCATTAAGAGTATGTTGTTGATATGGAATATTTACTGTGGGAATC 2700
2701 TTCTCACTGGTGTGAATCTTACCTGGCATTCGCGTTGATGCTAACTTCTACAACTG 2760
2701 TTCTCACTGGTGTGAATCTTACCTGGCATTCGCGTTGATGCTAACTTCTACAACTG 2760

2761 ATTCAAAATGGATTTAAAAATGATCAGCCATTTTATGCTACAGAAATATACATTATA 2820
2761 ATTCAAAATGGATTTAAAAATGATCAGCCATTTTATGCTACAGAAATATACATTATA 2820
2821 ATGCAATCCCTGCTGGGCTTTTACTCAAGAAAGCGGCATCTTCCCTTAATTTGACTTCG 2880
2821 ATGCAATCCCTGCTGGGCTTTTACTCAAGAAAGCGGCATCTTCCCTTAATTTGACTTCG 2880
2881 TTTTAGGATGTACGTGGGATGCGAAGAGGATGATACAGATGTGATGATGATGATGATG 2940
2881 TTTTAGGATGTACGTGGGATGCGAAGAGGATGATACAGATGTGATGATGATGATGATG 2940
2941 GTTTCGAATGTCTCACACCTTACCAAAACAGGCGCTTTTACAGAGAGATGATGATG 3000
2941 GTTTCGAATGTCTCACACCTTACCAAAACAGGCGCTTTTACAGAGAGATGATGATG 3000
3001 GGGTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
3001 GGGTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGAGCTTTTCTCTAGAAGCCGCT 3180
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGAGCTTTTCTCTAGAAGCCGCT 3180
3181 GCGTTACTCTGTTTCAAAGGACTTTTGTAAATCAAAATCAATCTCTCAGAGGCG 3240
3181 GCGTTACTCTGTTTCAAAGGACTTTTGTAAATCAAAATCAATCTCTCAGAGGCG 3240
3241 GAGAGCTGATATGAATTTTATGAGCATGTGATCTGATCAAGGCTTCTCAGGCG 3300
3241 GAGAGCTGATATGAATTTTATGAGCATGTGATCTGATCAAGGCTTCTCAGGCG 3300
3301 GCTTGAAGTATGCTGACCTGAGTACAGTATATCTTGTAAATCAATCAAAACAAG 3360
3301 GCTTGAAGTATGCTGACCTGAGTACAGTATATCTTGTAAATCAATCAAAACAAG 3360
3361 ATTTGTAAGGAGCACTAATATGATTTTAAAGTCTATGTTTAAATATATGATAA 3420
3361 ATTTGTAAGGAGCACTAATATGATTTTAAAGTCTATGTTTAAATATATGATAA 3420
3421 TTTTTCAGCTATTTAGTATATATTTTATGAGTGGGATATAAATTTCTACTACAGAAA 3480
3421 TTTTTCAGCTATTTAGTATATATTTTATGAGTGGGATATAAATTTCTACTACAGAAA 3480
3481 AAAAAAAAAAAAAAAAAAAAAA 3501
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 2

US-09-872-136-3

; Sequence 3, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; STREET: ImClone Systems Incorporated

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/872,136

FILING DATE: 01-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786

FILING DATE: <Unknown>

APPLICATION NUMBER: US/09/021,324

FILING DATE: <Unknown>

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEW-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3501 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 58..3039

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 139..3036

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-872-136-3

Query Match 100.0%; Score 3501; DB 10; Length 3501;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

61 CGCGCGTGGCGCGGACCGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 120

61 CGCGCGTGGCGCGGACCGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 120

121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAAG 180

121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAAG 180

181 AACAAATGATTCATCAGTGGGAAGTCATCATATCCATGGTATCAGAATCCCGGAA 240
181 AACAAATGATTCATCAGTGGGAAGTCATCATATCCATGGTATCAGAATCCCGGAA 240
241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTCCCGCTGTG 300
241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTCCCGCTGTG 300
301 GAAGTGGATGTATCTGCTTCCATCATACATGCAAGTGTGCTGATGCCCGGGAACATT 360
301 GAAGTGGATGTATCTGCTTCCATCATACATGCAAGTGTGCTGATGCCCGGGAACATT 360
361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAAATGCGACGACCATTTGATTTACAA 420
361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAAATGCGACGACCATTTGATTTACAA 420
421 AACAGGAGGATTTGTTTCCATGCTCATTTTGAATAACACAAACCCCAAGCTGGAATAC 480
421 AACAGGAGGATTTGTTTCCATGCTCATTTTGAATAACACAAACCCCAAGCTGGAATAC 480
481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTGTTACAGTGAATTAAGA 540
481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTGTTACAGTGAATTAAGA 540
541 AATACCTCTCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGAGCC 600
541 AATACCTCTCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGGAGCC 600
601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTTCGATTTCA 660
601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTTCGATTTCA 660
661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720
661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720
721 CATGAATTTATTCGGAGGACATTAAGTGTCTGCTGCCAGAAATGAACGAGGAGTGC 780
721 CATGAATTTATTCGGAGGACATTAAGTGTCTGCTGCCAGAAATGAACGAGGAGTGC 780
781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCCACATTTGCCACAATTTT 840
781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCCACATTTGCCACAATTTT 840
841 CTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTTCAATGTAACCATGGATTC 900
841 CTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTTCAATGTAACCATGGATTC 900
901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCAGGAGGCGCACTACTTTGAGATGAGT 960
901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCAGGAGGCGCACTACTTTGAGATGAGT 960
961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTCTGTTTGTATCATCAGTGGA 1020
961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTCTGTTTGTATCATCAGTGGA 1020
1021 AGAAACGACACCGGATATACACTTGTCTCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1021 AGAAACGACACCGGATATACACTTGTCTCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATATGAAATT 1140
1081 GTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATATGAAATT 1140
1141 GACCAATATGAAGATTTTGTCTCTGCTAGGTTTAAAGCTACCCACAATCAGATGT 1200
1141 GACCAATATGAAGATTTTGTCTCTGCTAGGTTTAAAGCTACCCACAATCAGATGT 1200
1201 ACCTGGAGCTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATTAACGATAC 1260
1201 ACCTGGAGCTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATTAACGATAC 1260
1261 AGCATATCAAGTTTGTGAATCATTAAGCACGAGGAGGAGATATATATTCATGCAGAA 1320

1261 1261 AGCATATCCAAGTTTTCATATCAAGACACAGCCAGGAGAAATATATTCATTCAGAGAA 1320
1321 1321 AATGATGATGCCAATTTACAAAATGTTTCAGCGTGAATATATAAGAGAAACCTCAAGTG 1380
1321 1321 AATGATGATGCCAATTTACAAAATGTTTCAGCGTGAATATATAAGAGAAACCTCAAGTG 1380
1381 1381 CTCGCAGAACATCGGCAAGTCAAGCGTCTGTTTCTCGATGATACCCATACCACATCT 1440
1381 1381 CTCGCAGAACATCGGCAAGTCAAGCGTCTGTTTCTCGATGATACCCATACCACATCT 1440
1441 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCGACAGAGAGATCACAGAAAGGA 1500
1441 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCGACAGAGAGATCACAGAAAGGA 1500
1501 1501 GTCTGGAATAGAAGGCTACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACTCTA 1560
1501 1501 GTCTGGAATAGAAGGCTACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACTCTA 1560
1561 1561 AACATGATGAAGCCATAAAGGTTCTCGTCAAGTGCCTGATACAAATTCCTTTGGC 1620
1561 1561 AACATGATGAAGCCATAAAGGTTCTCGTCAAGTGCCTGATACAAATTCCTTTGGC 1620
1621 1621 ACATCTGTGAGAGTCTTTTAAACTCTCCAGGCCCTTCCCTTTATCTCAAGACAAC 1680
1621 1621 ACATCTGTGAGAGTCTTTTAAACTCTCCAGGCCCTTCCCTTTATCTCAAGACAAC 1680
1681 1681 ATCTCATCTATGACAAATGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1681 1681 ATCTCATCTATGACAAATGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1741 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1741 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1801 1801 ACCGGTCTCTCAGATATGATGATCTCTACGTTGATTTACAGAGAAATGATATGATCTC 1860
1801 1801 ACCGGTCTCTCAGATATGATGATCTCTACGTTGATTTACAGAGAAATGATATGATCTC 1860
1861 1861 AATGGAGTTTCCAGAGAGAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1861 1861 AATGGAGTTTCCAGAGAGAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1921 1921 GGAAGATGATGAACCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 1921 GGAAGATGATGAACCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 1981 GCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
1981 1981 GCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
2041 2041 CTCAGATGATGCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGCGTGC 2100
2041 2041 CTCAGATGATGCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGCGTGC 2100
2101 2101 ACATGTCAGGACCAATTTACTTGAATTTGATGATGTTGCTATGTTGATCTCTCAAC 2160
2101 2101 ACATGTCAGGACCAATTTACTTGAATTTGATGATGTTGCTATGTTGATCTCTCAAC 2160
2161 2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTTCAGAGAAACAC 2220
2161 2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTTCAGAGAAACAC 2220
2221 2221 AATTTTCAGTTTACCCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGTTCAAGA 2280
2221 2221 AATTTTCAGTTTACCCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGTTCAAGA 2280
2281 2281 GAAGTTTCAGATACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340
2281 2281 GAAGTTTCAGATACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340
2341 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGGAGGACTTGAAT 2400
1261 1261 AGCATATCCAAGTTTTCATATCAAGACACAGCCAGGAGAAATATATTCATTCAGAGAA 1320
1321 1321 AATGATGATGCCAATTTACAAAATGTTTCAGCGTGAATATATAAGAGAAACCTCAAGTG 1380
1321 1321 AATGATGATGCCAATTTACAAAATGTTTCAGCGTGAATATATAAGAGAAACCTCAAGTG 1380
1381 1381 CTCGCAGAACATCGGCAAGTCAAGCGTCTGTTTCTCGATGATACCCATACCACATCT 1440
1381 1381 CTCGCAGAACATCGGCAAGTCAAGCGTCTGTTTCTCGATGATACCCATACCACATCT 1440
1441 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCGACAGAGAGATCACAGAAAGGA 1500
1441 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCGACAGAGAGATCACAGAAAGGA 1500
1501 1501 GTCTGGAATAGAAGGCTACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACTCTA 1560
1501 1501 GTCTGGAATAGAAGGCTACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACTCTA 1560
1561 1561 AACATGATGAAGCCATAAAGGTTCTCGTCAAGTGCCTGATACAAATTCCTTTGGC 1620
1561 1561 AACATGATGAAGCCATAAAGGTTCTCGTCAAGTGCCTGATACAAATTCCTTTGGC 1620
1621 1621 ACATCTGTGAGAGTCTTTTAAACTCTCCAGGCCCTTCCCTTTATCTCAAGACAAC 1680
1621 1621 ACATCTGTGAGAGTCTTTTAAACTCTCCAGGCCCTTCCCTTTATCTCAAGACAAC 1680
1681 1681 ATCTCATCTATGACAAATGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1681 1681 ATCTCATCTATGACAAATGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1741 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1741 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
1801 1801 ACCGGTCTCTCAGATATGATGATCTCTACGTTGATTTACAGAGAAATGATATGATCTC 1860
1801 1801 ACCGGTCTCTCAGATATGATGATCTCTACGTTGATTTACAGAGAAATGATATGATCTC 1860
1861 1861 AATGGAGTTTCCAGAGAGAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1861 1861 AATGGAGTTTCCAGAGAGAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
1921 1921 GGAAGATGATGAACCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 1921 GGAAGATGATGAACCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 1981 GCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
1981 1981 GCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
2041 2041 CTCAGATGATGCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGCGTGC 2100
2041 2041 CTCAGATGATGCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGGCGTGC 2100
2101 2101 ACATGTCAGGACCAATTTACTTGAATTTGATGATGTTGCTATGTTGATCTCTCAAC 2160
2101 2101 ACATGTCAGGACCAATTTACTTGAATTTGATGATGTTGCTATGTTGATCTCTCAAC 2160
2161 2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTTCAGAGAAACAC 2220
2161 2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTTCAGAGAAACAC 2220
2221 2221 AATTTTCAGTTTACCCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGTTCAAGA 2280
2221 2221 AATTTTCAGTTTACCCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGTTCAAGA 2280
2281 2281 GAAGTTTCAGATACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340
2281 2281 GAAGTTTCAGATACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340
2341 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGGAGGACTTGAAT 2400

2341 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
2401 2401 GTGCTTACATTTGAAGATCTTCTTGTTCGATATCAAGTTGCGCAAGGAATGAATTT 2460
2401 2401 GTGCTTACATTTGAAGATCTTCTTGTTCGATATCAAGTTGCGCAAGGAATGAATTT 2460
2461 2461 CTGGAATTTAAGTCGTTGTTTACAGACACCTGGCCGCGAGAACTGCTTCTCAGCCAC 2520
2461 2461 CTGGAATTTAAGTCGTTGTTTACAGACACCTGGCCGCGAGAACTGCTTCTCAGCCAC 2520
2521 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGATGATTTCCAAC 2580
2521 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGATGATTTCCAAC 2580
2581 2581 TATGTTGTGAGGGCAATGCGCTGCTGTAAATGGATGGCCCCGAGAAACGCTGTTT 2640
2581 2581 TATGTTGTGAGGGCAATGCGCTGCTGTAAATGGATGGCCCCGAGAAACGCTGTTT 2640
2641 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTTCATATGGAATATTACTGTGGAAATC 2700
2641 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTTCATATGGAATATTACTGTGGAAATC 2700
2701 2701 TTCTCAGTTGTTGAATCTTACCTGGCATTTCCGTTTGTGCTTAACTTCTACAAACG 2760
2701 2701 TTCTCAGTTGTTGAATCTTACCTGGCATTTCCGTTTGTGCTTAACTTCTACAAACG 2760
2761 2761 ATTCAAATGATTTTAAATGGATCAGCCATTTTATGTACAGAGAAATATACATATA 2820
2761 2761 ATTCAAATGATTTTAAATGGATCAGCCATTTTATGTACAGAGAAATATACATATA 2820
2821 2821 ATGCAATCTCTGCGCTTTTGAAGCAAGGCAAGGCGCATCTTCCCTTAAATTTGACTTCG 2880
2821 2821 ATGCAATCTCTGCGCTTTTGAAGCAAGGCAAGGCGCATCTTCCCTTAAATTTGACTTCG 2880
2881 2881 TTTTATGATGTCAGTGGCAGATGCAGAAAGGAGTATCATCAAGATGATGATGCGCT 2940
2881 2881 TTTTATGATGTCAGTGGCAGATGCAGAAAGGAGTATCATCAAGATGATGATGCGCT 2940
2941 2941 GTTTCGGAATGCTCTCACACCTTACCAAGGCGACCTTTTACAGAGAGATGATTTG 3000
2941 2941 GTTTCGGAATGCTCTCACACCTTACCAAGGCGACCTTTTACAGAGAGATGATTTG 3000
3001 3001 GGGTACTCTCTCGCAGGCTCAGTCCAAAGATGCTAGAGAGAAATTTAGTTTAAAG 3060
3001 3001 GGGTACTCTCTCGCAGGCTCAGTCCAAAGATGCTAGAGAGAAATTTAGTTTAAAG 3060
3061 3061 ACTTCATCTCTCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTCA 3120
3061 3061 ACTTCATCTCTCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTCA 3120
3121 3121 CACTAAAAGAAATCTATTATCAACTGCTTCCACAGACTTTTCTCTAGAGCGCTCT 3180
3121 3121 CACTAAAAGAAATCTATTATCAACTGCTTCCACAGACTTTTCTCTAGAGCGCTCT 3180
3181 3181 GCGTTTACTCTGTTTCAAAGGAGCTTTTGTAAATCAATCATCTCTGTCACAGGAG 3240
3181 3181 GCGTTTACTCTGTTTCAAAGGAGCTTTTGTAAATCAATCATCTCTGTCACAGGAG 3240
3241 3241 GAGGAGCTGATAATGAACCTTTTATGGAGCTGATCTCATCCAGGCGCTTCTCAGGCG 3300
3241 3241 GAGGAGCTGATAATGAACCTTTTATGGAGCTGATCTCATCCAGGCGCTTCTCAGGCG 3300
3301 3301 GCTTGAAGTGAATCTGTACCTTGAAGTACAGTATATTTCTGTAATACATAAAACAAAGC 3360
3301 3301 GCTTGAAGTGAATCTGTACCTTGAAGTACAGTATATTTCTGTAATACATAAAACAAAGC 3360
3361 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATCTTTTAAATATATATGTA 3420
3361 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATCTTTTAAATATATATGTA 3420
3421 3421 TTTTTCAGCTATTTAGTATATATTTATGGTGGGAATAAATTTCTTACTACAGAAAA 3480
3421 3421 TTTTTCAGCTATTTAGTATATATTTATGGTGGGAATAAATTTCTTACTACAGAAAA 3480

QY	1318	GA	AATGATGATGCGCCAAATTTACCAAAATGTTTCACGCTGGAATATAAGAAAGGAACCTCAA	1317
DB	1261	GA	AAATGATGATGCGCCAAATTTACCAAAATGTTTCACGCTGATATAAAGAAAGGAACCTCAA	1320
QY	1378	GT	GCTCGCAGAAAGCATGGCAAGTCAGCGTCCTGTTTCTCGGATGATACCCATTACCA	1437
DB	1321	GT	CTCGCAGAAAGCTTCGGCAAGTCAGCGTCCTGTTTCTCGGATGATACCCATTACCA	1380
QY	1438	TC	TTTGGACCTGGAAGAAGTTTCAGACAAGTCTCCCAACTGCCACAGAAGAGATCACAGAA	1497
DB	1381	TC	TTTGGACCTGGAAGAAGTTTCAGACAAGTCTCCCAACTGCCACAGAAGAGATCACAGAA	1440
QY	1498	GG	AGTCTGGAATAGAAAGGCTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACT	1557
DB	1441	GG	AGTCTGGAATAGAAAGGCTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACT	1500
QY	1558	CT	AAACATGAGTGAAGGCATAAAAGGTTCTCGTCAAGTCTGCTGCATACAATTTCCCTT	1617
DB	1501	CT	AAACATGAGTGAAGGCATAAAAGGTTCTCGTCAAGTCTGCTGCATACAATTTCCCTT	1560
QY	1618	GG	CACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTCATCTCCAAGAC	1677
DB	1561	GG	CACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTCATCTCCAAGAC	1620
QY	1678	AA	CATCTATTCATGCAACAATTTGGTGTGTTGCTCTCTCTTCATGTCGTTTAAACCTG	1737
DB	1621	AA	CATCTATTCATGCAACAATTTGGTGTGTTGCTCTCTCTTCATGTCGTTTAAACCTG	1680
QY	1738	CT	AATTTGTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
DB	1681	CT	AATTTGTACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
QY	1798	GT	CACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATGAT	1857
DB	1741	GT	CACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATGAT	1800
QY	1858	CT	CAATGGGAGTTTCCAAGAGAAAATTTAGATTTGGGAAGGTACTTAGGATCAGGTGCT	1917
DB	1801	GT	CAATGGGAGTTTCCAAGAGAAAATTTAGATTTGGGAAGGTACTTAGGATCAGGTGCT	1860
QY	1918	TT	TGAAAAGTGATGAACGCAACAGCTTATGGAATAGCAAAACAGAGTCTCAATCCAG	1977
DB	1861	TT	TGAAAAGTGATGAACGCAACAGCTTATGGAATAGCAAAACAGAGTCTCAATCCAG	1920
QY	1978	GT	TGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTCA	2037
DB	1921	GT	TGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTCA	1980
QY	2038	GA	ACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGGG	2097
DB	1981	GA	ACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGGG	2040
QY	2098	TG	CACACTGTCAGGACCAATTTACTGATTTTGAATCTGTTGCTGATGCTCTCTC	2157
DB	2041	TG	CACACTGTCAGGACCAATTTACTGATTTTGAATCTGTTGCTGATGCTCTCTC	2100
QY	2158	AA	CATCTAAGAAGTAAAAGAGAAAATTTTCAGGACTTGGACAGAGATTTTTCAGAGAA	2217
DB	2101	AA	CATCTAAGAAGTAAAAGAGAAAATTTTCAGGACTTGGACAGAGATTTTTCAGAGAA	2160
QY	2218	CA	CAATTCAGTTTTTACCCCACTTTCCAAATACATCCAAATPCCAGCATGCTGTCTCA	2277
DB	2161	CA	CAATTCAGTTTTTACCCCACTTTCCAAATACATCCAAATPCCAGCATGCTGTCTCA	2220
QY	2278	AG	AGAAAGTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTT	2337
DB	2221	AG	AGAAAGTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTT	2280
QY	2338	CA	CTCTGAAGATCAAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGGAGGACTTGG	2397
DB	2281	CA	CTCTGAAGATCAAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGGAGGACTTGG	2340
QY	2398	AA	TGCTTACATTTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAA	2457

RESULT 4

```

US-09-919-408-1
; Sequence 1, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
;
; APPLICANT: Lemischka, Ihor R.
;
; TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

```

248	AAAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATCACCTCC	307
332	AAGTGTGGTTCGATGCGCCAGGGAACAATTTCTGTCTCTGGGTCTTTAAAGACAGCTCCC	391
308	AAGTGCAGCTCGCCACCCAGGGGACCTTTCTCTGGGTCTTTAAAGACAGCTCCC	367
392	TGAATTTGCCAGCCACATTTTGAATTTACAAACAGAGAGGTGTTTCCATGGTTCATTTGA	451
368	TGGCTGCCAGCCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATTTGA	427
452	AAATGACAGAAACCCAGCTGGGAAATACCTACTCTTTTATTCAGAGTGAAGCTACCAAT	511
428	ACGTGACAGAGACCCAGCGAGGAAATACCTTACTTCCATATTTACAGAGCAACGCGCACT	487
512	ACACAATATTCTTTACAGTGAATATAAGAAATACCTCTCTTTACACATTTAAGAGACCTT	571
488	ACACAGTACTGTTTACAGTGAATGTAGAGATACACAGCTGTATGTCTAAGGAGACCTT	547
572	ACTTTAGAAAAATGAAAAACAGGACCCCTGGTCTCGCATATTCAGAGCGTTTCCAGAGC	631
548	ACTTTAGGAAGATGAAAAACAGGATGCACTGCTCTGCATCTCCGAGGGTGTTCGCGAGC	607
632	CGATCGTGAATGGGTGCTTTGCGATTCACAGSGGGGAAGCTGTAAAGAGAAAGTCCAG	691
608	CCACTGTGGAGTGGTCTCTGCAGCTCCCAAGGAAAGCTGTAAAGAGAAAGGCGCTG	667
692	CTGTTTAAAGAGGAGAAAGTCTTCATGATTAATTTGGGACGAGACATAAGGTGCT	751
668	CTGTTGTCAGAAAGGAGAAAGTACTTTCATGAGTCTTGGACACAGCATCAGATGCT	727
752	GTGCCAGAAATGAATGGGCGAGGAAATGACAGGCTGTTCCACAAATAGATCTAAATCAA	811
728	GTGCTAGAAATGCATGGGCCGGAATGCACCAAGCTGTTCCACCATAGATCTAAACCAAG	787
812	CTCCTCAGACACATGCGCACAAATTAATTTCTTAAAGTAGGGGAACCCCTTATGGATAAGT	871
788	CTCCTCAGAGCACACTGCCCCAGTTATTTCTGAAAGTGGGGAAACCCCTTGTGGATCAGGT	847
872	GCAAGCTGTTTATGTGACCATGGATTCGGGCTCACCTGGGAAATAGAAAAACAAGCAC	931
848	GTAAAGCCATCCATGTGAACCATGGATTCGGCTCACCTGGGAGCTGGAAGCAACAGCC	907
932	TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA	991
908	TGAGGAGGCGACCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA	967
992	TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAAGACACCGGATACCTACACTTTCTCT	1051
968	TTCTTTGGCTTTTGTGCTTCCGTGGGAAAGCAACGACACCGGATATTACACCTCTCTT	1021
1052	CTTCAAGATCCCACTCAATCAGCTTTTGGTTTACCATCGTAGGAAGGATTTATAAATG	1111
1028	CCTCAAGACACCCAGCCAGTCCAGCGTTTGGTGACCATCTAGAAAAGGTTTATAACG	1081
1112	CTACCAATTCAGTGAAGATTTGAAATTTGACCAATATGAAGAGTTTGTCTTTCTGTCA	1171
1088	CTACCACTCGAAGAAAGATGAAATTTGACCCGTACGAAAGTTCTCTCTCAGTCA	1141
1172	GGTTTAAAGCCTACCCACAAATCAGATGTACGTGGACCTTCTCTCGAAATCATTTCTCT	1231
1148	GGTTTAAAGCGTACCCAGGAATCCGATGACGTGGATCTTCTCTCAAGCTCATTTCTCT	1201
1232	GTGAGCAAAAGGCTTTGATAACGGATACAGATATCCAAGTTTTCATCATCAAGCAC	1291
1208	GTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTCGATCATAGAACA	1261
1292	AGCAGAGAGATATATATTCATGCGAGAAATGATGATGCCAATTTTACCAAAATGTTCA	1351
1268	AGCCAGAGAGATACATATTTCTATGCGAGAAATGATGACGCCAGTTCCACCAATGTTCA	1321
1352	CGCTGAATATAAGAAAGAAACCTCAAGTCTCGCAGAAAGATCGGAAAGTACGAGCGCTCT	1411

Db 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAGCCAGCGCTCCT 1387
QY 1412 GTTTCCTCGGATGATACCCATTACCATTCTTGGACCTTGGGAAGAAGTGTTCACAGAAGTCTC 1471
Db 1388 GTTCTCTGATGGCTACCCGCTACCCTCTTGGACCTTGGGAAGAAGTGTTCGGAACAATCTC 1447
QY 1472 CCAACTGCACAGAGAAGATCACAGAAGGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT 1531
Db 1448 CCAATTGCACGGAGAAATCCCAAGAAGGAGTTTGGAAATAAAAGGCTAACAGAAAAGTGT 1507
QY 1532 TTGGACAGTGGGTGTCAGACAGTACTCTTAAACATGAGTGAAGCCATAAAGAGGTTCCTGG 1591
Db 1508 TTGGCCAGTGGGTGTCAGACAGTACTCTTAAATATGAGTGGGCCGGAAGAGGCTTCCTGG 1567
QY 1592 TCAAGTGTGTGCATACAATTCCTTGGCACATCTTGTGAGACGATCCCTTTTAAACCTC 1651
Db 1568 TCAATGTCTGTGCTACAAATCTATGGGCACGCTCTTGGCAACCATCTTTTAAACCTAC 1627
QY 1652 CAGGCCCTTCCTTTTCATCCAAGACAACATCTCATCTATGCAACAATTTGGTGTGTC 1711
Db 1628 CAGGCCCTTCCTTTTCATCCAAGACAACATCTCTTCTATGCGACCATTTGGGCTCTGTC 1687
QY 1712 TCCTCTTCATTTGCTTTTAAACCTGCTAATTTGTCTACAGTACAAAAGCAATTTAGGT 1771
Db 1688 TCCCTTCTAATTTGTCTCATTTGTGTGATCTGCCACAAATACAAAAGCAATTTAGGT 1747
QY 1772 ATGAAAGCCAGCTACAGTGGTACAGTGCACCGCTCCTCAGATATGAGTACTCTTACG 1831
Db 1748 AGCAGAGTCAGCTGCAGATGATCCAGTGACTGCCCCCTGGATAACGAGTACTCTTACG 1807
QY 1832 TTGATTTTCAGAAATATGAATATGATCTCAAAATGGGAGTTTCCAAAGAGAAAATTTAGAGT 1891
Db 1808 TTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAACTTAGAGT 1867
QY 1892 TTGGGAGTACTAGGATCAGTGTCTTTTGGAAAAGTGAAGACGCAACAGCTTATGGAA 1951
Db 1868 TTGGGAAAGTCTGGGGTCTGGCGCTTTCGGGAGGATGATGACGCGCACCGCTATGGCA 1927
QY 1952 TTAGCAAAACAGGAGTCTCAATCCAGTTTGCCTCAAAATGCTGAAAGAAAAGCAGACA 2011
Db 1928 TTAGTAAACGGAGTCTCAATTCAGTGGCGTGAAGTCTTAAAGAGAAAGCTGACA 1987
QY 2012 GCTCTGAAAGAGGACACTCATGTCAAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
Db 1988 GCTGTGAAAAGAGGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
QY 2072 AGAATATGTGAACCTGTGGGGCTGCACACTGTACAGACCAATTTACTTGATTTTGG 2131
Db 2048 ACAACATCTGAATCTGCTGGGGCATGCACACTGTACAGGGCCAGTGTACTTGATTTTGG 2107
QY 2132 AATACTGTTGCTATGGTGTATCTTCAACTATCTAAGAAGTAAAGAGAAAATTTTCACA 2191
Db 2108 AATATTTGCTATGGTGACCTCTCACTTAAAGAAAGTAAAGAGAGAGTTTCACA 2167
QY 2192 GGACTTGGACAGAGATTTTCAAGGAACAAATTTACAGTTTAAACCCACTTTCCTCAATCAC 2251
Db 2168 GGACATGGACAGAGATTTTAAAGAACATAATTTACAGTTCTTACCCTACTTCCAGGCAC 2227
QY 2252 ATCCAAATTCACAGATGCTGTGTTCAAGAGAGTTCAGATACACCCGGACTCGGATCAAA 2311
Db 2228 ATTCAAATTCACAGATGCTGTGTTACAGGAGAGTTTCAGTTACACCCGCTTGGATCAGC 2287
QY 2312 TCTCAGGGCTTCATGGGAATTCATTTCACTCTCAGATGAAATGGAATATGAACCAAA 2371
Db 2288 TCTCAGGGTTCATGGGAATTCATTTCTTCAAGATGAGATGGAATATGAACCAAA 2347
QY 2372 AAAGGCTG-----GAAAGAGGAGGAGCTTGAATGTGCTTACATTTGAAGATCTTCTTT 2425
Db 2348 AGAGGCTGCCAAGAGAGGAGGAGAGATTTGACGTGCTGACGTTTGAAGACCTCCTTT 2407
QY 2426 GCTTTGCATATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTCACA 2485
Db 2408 GCTTTGCGTACCAAGTGGCCAAAGGATCGAATTCCTGGAGTTCGAAGTCTGTGTGCCACA 2467

QY 2486 GAGACCTGGCCGCCAGGAAGCTGCTTCTCACCCACGGGAAGTGGTGAAGATATGTGACT 2545
Db 2468 GAGACCTGGCAGCAGGAATGTGTGTGTACCCACGGGAAGGTGGTGAAGATCTGTGACT 2527
QY 2546 TTGGATTGGCTCGAGATATCATGAGTATTCCAACTATGTGTTCAGGGCAATGCCGCTC 2605
Db 2528 TTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTAGCTCGTCAGGGCAACCCACGGC 2587
QY 2606 TGCTGTAAATGGATGGCCCCCGAAAGCCTGTTTGAAGGCATCTACACCAATTAAGAGTG 2665
Db 2588 TGCCGCTGAAGTGGATGGCACCAGAGACTTATTGAAGGGATCTACACAATCAAGAGTG 2647
QY 2666 ATGCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAGTGTGGTGTGAATCCTTACC 2725
Db 2648 ACCTGTGGTCTACGGCATCCTTCTCTGGGAGATATTTTACCTGGGTGTGAACCTTACC 2707
QY 2726 CTGGCATTCGGTGTGATGCTTAACCTTCTACAACTGATTCAAAATGGATTTAAAATGGATC 2785
Db 2708 CTGGCATTCCTGTGCGAGCTAACTTCTATAAATGATTACAGAGTGGATTTAAAATGGAGC 2767
QY 2786 AGCCATTTTATGCTACAGAGAATAATATACATTAATATGCAATCCTGCTGGGCTTTTGACT 2845
Db 2768 AGCCATTTCTATGCCACAGAGGATATACTTTGTATATGCAATCCTGCTGGGCTTTTGACT 2827
QY 2846 CAAGGAAACGCCATCCTCCCTAAATTTGACTTCTGTTTTTAGGATGTACAGTGCAGATG 2905
Db 2828 CAAGGAACGGCCATCCTCCCAACCTGACTTCAATTTTAGGATGTACAGTGCAGAGG 2887
QY 2906 CAGAAGAAGCGATGTATCAGAATGTGATGCCGCTGTTTCGGATGTCTCCACACCTTACC 2965
Db 2888 CAGAAGAAG-----CATGTATCAGAACATCCATCTTACC 2924
QY 2966 AAAACAGCGACCTTTACAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGAGCTCAGG 3025
Db 2925 AAAACAGCGCGCCCTCAGCAGAGAGCGGCTCAGAGC--CCAGTCCACACGCGCAGG 2983
QY 3026 TCCAAAGATTCGTAGAGGAACAATTTAGTTTTAAGGACTTCACTCCCTCCACCTATCCCTAA 3085
Db 2984 T-GAAGATTCACAGAGAAAAGTTAGCGAGGAGGCTTGGACCCCGC----CACCCCTAG 3038
QY 3086 CAGGCTGTAGATTACCAAAAACAAGATTAAATTCATCACTAAAGAAAATCTATTATCAAC 3145
Db 3039 CAGGCTGTAGACCCGACAGCCAGATTAAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
QY 3146 TGCTGCTTACACAGACTTTTCTCTAGAAGCGCTGCTGTTTACTTGTGTTTTCAAAGGA 3205
Db 3097 CGTTGCTTGCCTGGACTTTTCTCTAGATGCTGTGCTGCTTACTC-----CAAAGTGA 3149
QY 3206 CTTTTCATAAATCAATCATCCTGTCAAGAGCAGGAGGAGCTGATTAATGAACCTTTATTG 3265
Db 3150 CTTCTATAAATCAAACTCTCCTCGCACAGCGGGAGAGCCAAATAATAGACTTTGTTGG 3209
QY 3266 GAGCATTTGATCTCATCCAAAGGCTTCTCAGGCGGCTTCAGTGAATTTGTACTCTCAAG 3325
Db 3210 TGAGCCCGCTACCTCGGGGCTTTCCAGCAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269
QY 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTGAAGAGAGC 3378
Db 3270 TATAGTATATCTTGTAAATACGTGAACAAACAAACCCGCTTTTGTGTAAGGAAGC 3329
QY 3379 TAATATGATTTTTT--AAGTCTATGTTTAAATATATATGTAATTTTTCAGCTATTATTAG 3436
Db 3330 TAAATATGATTTTTTAAATAATCTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
QY 3437 TGATATATTTTATGGGTGGAATAAATAATTTCTACTACAGAAAATAAAAAAATAAAAAA 3496
Db 3390 TGATATATTTTATGGATGGAATAAATAATTTCTACTGTAAAAATAAAAAAATAAAAAA 3449
QY 3497 AAAA 3500
Db 3450 AAAA 3453

```
NAME/KEY: CDS
LOCATION: 31...3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-872-136-1

Query Match
Best Local Similarity 63.6%; Score 2227.2; DB 10; Length 3453;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

35 GGGACCCCGGCTCCGGAGGCCATGCGCGCTTGGC---GCGCAGCGGGCAGCCGCTGC 91
8 GGCTACCGCGCTCCGGAGGCCATGCGCGCTTGGCAGCGCCAGCCGCGGCTGC 67
92 CGCTGCTGTTGTTTCTGCAATATATTGGGACTATTACAAATCAAGATCTCCCTG 151
68 TGCTGCTGTTGTTTCTGCAATATATTGGGACTATTACAAATCAAGATCTCCCTG 127
152 TGATCAAGTGTGTTTAAATCAATCAATAGAAATGATTATCATCAGTGGGAGTCAAT 211
128 TGATCAAGTGTGTTTAAATCAATCAATAGAAATGATTATCATCAGTGGGAGTCAAT 187
212 CATATCCCATGTTATCAGAAATCCCGGAGAGCTCGGGTGTGGTGTGAGACCCAGACT 271
188 CGTACCGAATGTTGCGAGGATCCCGAAGAGCTCCAGTGTACCCGAGCGCCAGAGTG 247
272 CAGGACAGTGTACGAAGCTGCGGCTGTGGAAGTGTATCTGCTCCATCAGACTGC 331
248 AAGGACGCTATATGAAGCGGCGCCGCTGGAGGTGGCGGAGTCTGGTCCATCACCCTGC 307
332 AAGTGTGCTGATGCTCCCGGAGAGCTTCTGCTCTGGGTCTTTAAGCAGACTGCC 391
308 AAGTGTGCTGCTGCTCCCGGAGAGCTTCTGCTCTGGGTCTTTAAGCAGACTGCC 367
392 TGAATTCAGGACCATTTTGAATTTACAAACAGAGAGTGTGTTCCATGGTCAATTTGA 451
368 TGGGCTGCGGCGGCTTTGATTTACAAACAGAGAGTGTGTTCCATGGTCAATTTGA 427
452 AAATGACAGAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGTCAATTTGA 511
428 ACCTGACAGAGACCCAGGAGAGATACCTACTTCCATATTCAGAGGAGAGCGGCAACT 487
512 ACACATATTTGTTACAGTGTATAGAAATACCCCTGCTTTACACATTAAGAGAGCTT 571
488 ACACATATTTGTTACAGTGTATAGAAATACCCCTGCTTTACACATTAAGAGAGCTT 547
572 ACTTTAGAAAAATGAAACAGGAGCGCTTGTGCTCATATCTGAGAGGCTTCCAGAGC 631
548 ACTTTAGAGAGATGAAACAGGAGCTGCTGCTCATCTCCGAGGGTGTTCGGAGC 607
632 CGATCTGGAATGGTGTGTTGATTCAGAGGAGAGAGCTGTAAAGAGAGAGTCCAG 691
608 CCACCTGGAGTGGTGTCTGCTGAGCTCCACAGGAGAGAGCTGTAAAGAGAGAGTCCAG 667
692 CTGTTGTTAAAGAGAGAGAGAGTGTCTCATGATTTTGGGAGAGAGAGAGTGTCT 751
668 CTGTTGTTAAAGAGAGAGAGAGTGTCTCATGATTTTGGGAGAGAGAGAGTGTCT 727
752 GTCCAGAAATGAATGGGAGGAGATGCACAGGCTGTTTCAACAATAGATTAATCAAA 811
728 GTCTAGAAATGAATGGGAGGAGATGCACAGGCTGTTTCAACAATAGATTAATCAAA 787
812 CTCTCAGACCATTCATGCTGCAATTTTCTTAAAGTAGGGAGAGCTTATGGATAGGT 871
788 CTCTCAGACCATTCATGCTGCAATTTTCTTAAAGTAGGGAGAGCTTATGGATAGGT 847
872 GCAAGCTGTTTCAATGCAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931
848 GTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
932 TCGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
908 TGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967
```

Qy 992 TTCTGTTTGGCTTTTGTATCATAGTGGCAAGAACGACACCGGATACACTTGTTCCT 1051
Dy 998 TTCTCTGGCTTTGTCTCTCGTGGGAAGAACGACACCGGATATACACCTGCTCT 1027
Qy 1052 CTTCAAGACATCCAGTCAATACGCTTGTGTTACCATCGTAGGAAGAGGATTTATAATG 1111
Dy 1028 CTTCAAGACATCCAGTCAATACGCTTGTGTTACCATCGTAGGAAGAGGATTTATAACG 1087
Qy 1112 CTACCAATTCAGTGAAGATTTAGAAATTCACCAATATGAAGAGTCTTTGTTTCTCTCA 1171
Dy 1088 CTACCAAGTTCAGTGAAGATTTAGAAATTCACCAATATGAAGAGTCTTTGTTTCTCTCA 1147
Qy 1172 GGTTTAAAGCCTACCCACAAATCAGATGTACGTTGACCTTCTCTCGAAATATCAATTCCTT 1231
Dy 1148 GGTTTAAAGCCTACCCACAAATCAGATGTACGTTGACCTTCTCTCGAAATATCAATTCCTT 1207
Qy 1232 GTGAGCAAAAGGCTCTGTATAACGGATACAGATATCAAGTCTTTGCAATCATAGACAC 1291
Dy 1208 GTGAACAGAGAGCCCTGGAGATGGGTACAGATATCTAAATTTTGGGATCATAGAACA 1267
Qy 1292 AGCCAGGAGAAATATATATTCATCGAGAAATGATGATGCCCAATTTACCAAAATGTCA 1351
Dy 1268 AGCCAGGAGATACATATCTATGACAGAAATGATGATGCCCAATTTACCAAAATGTCA 1327
Qy 1352 CGCTGAATATAAGAGGAAACCTCAAGTGTCTGCAGAGATCGGCAAGTCAGCGCTCT 1411
Dy 1328 CGCTGAATATAAGAGGAAACCTCAAGTGTCTGCAGAAATGCTCAGCCAGCCAGCGCTCT 1387
Qy 1412 GTTCTCGATGATACCATTCATCTTGGACCTGGAGAGATGTTCAAGCAAGTCTC 1471
Dy 1388 GTTCTCTGATGCTACCCGCTACCTCTTGGACCTGGAGAGATGTTCAAGCAAGTCTC 1447
Qy 1472 CCAACTGCACAGAGATACAGAGAGGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT 1531
Dy 1448 CCAATTTGCACGAGGAAATCCAGAGGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT 1507
Qy 1532 TTGGACAGTGGGTGTCCAGACGACTCTAAACATGAGTGAAGCCATAAAGAGTTCCTGG 1591
Dy 1508 TTGGCCAGTGGGTGTCCAGACGACTCTAAATATGAGTGAAGCCAGGCAAGGCTTCTGG 1567
Qy 1592 TCAAGTGTGTGATACAAATTCCTTGGACATCTTGTGAGAGATGCTTTAAACTCTC 1651
Dy 1568 TCAATATGCTGTGGCTCAATCTTATGGCAGCTCTTGGCAAGCATCTTTTAAACTCAC 1627
Qy 1652 CAGGCCCTTCCCTTTCATCCAAAGCAACATCTCATCTATGCAACAATTTGGTGTGTC 1711
Dy 1628 CAGGCCCTTCCCTTTCATCCAAAGCAACATCTCTTATGCAACAATTTGGTGTGTC 1687
Qy 1712 TCCCTTTCATTTGCTGTTTTAACCTGCTAAATTTGTCACAAAGTACAAAAGCAATTTAGGT 1771
Dy 1688 TCCCTTTCATTTGCTGTTTTAACCTGCTAAATTTGTCACAAAGTACAAAAGCAATTTAGGT 1747
Qy 1772 ATGAAGCCAGCTACAGATGGTACAGTGTACCGCTCCTCAGATATGAGTACTTCTACG 1831
Dy 1748 ACAGAGTACAGTGTACAGATGATCCAGTGTACTGGCCCTTGGATACGACTTCTACG 1807
Qy 1832 TTGATTTACAGAAATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGT 1891
Dy 1808 TTGACTTCAGGAGTATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGT 1867
Qy 1892 TTGGGAGGTACTAGGATCAGTGTGCTTTTGGAAAAGTGTAGAACGCAACAGCTTTATGGAA 1951
Dy 1868 TTGGGAGGTCTGGGCTTGGGCTTTTGGGAGGTGTAGAACGCAACGCTTATGGCA 1927
Qy 1952 TTAGCAAAACAGGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAAAAGCAGACA 2011
Dy 1928 TTAGTAAAACGGAGTCTCAATTCAGTGTGGGCTGAAAGTCTTAAAGAGAAAAGCTGACA 1987
Qy 2012 GCTCTGAAGAGAGGACATGCTCAGAACTCAAGATGATGACCCAGCTGGGAGGCCACG 2071
Dy 1988 GCTGTGAAGAAAGAGCTCTCATGCTGGAGCTCAAAATGATGACCCAGCTGGGAGGCCACG 2047
Qy 2072 AGAATATTGTGAACCTGCTGGGGCGGTGCACATGTCAGGACCAATTTACTTGTATTTTG 2131

Dy 2048 ACAACATCGTGAATCTGCTGGGGCATGCACACTGTACAGGGCCAGTGTACTTGTATTTTG 2107
Qy 2132 AATACTGTGCTATGCTGATCTTCTCAACTATCTAAGAACTAAAGAGAAAATTTTCACA 2191
Dy 2108 AATATTGTGCTATGCTGATCTTCTCAACTATCTAAGAACTAAAGAGAAAATTTTCACA 2167
Qy 2192 GGACTGGACAGAGATTTTCAAGGAACAATTTCAAGTCTTACCCACTTTCCAAATCAC 2251
Dy 2168 GGACATGGACAGAGATTTTAAAGAACTAATTTCAAGTCTTACCCACTTTCCAGGCAC 2227
Qy 2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAACTCAGATACACCCGACTCGGATCAAA 2311
Dy 2228 ATTCAAATTCAGCATGCTGTTTCAAGAGAACTCAGTACACCCGCTTGGATCAGC 2287
Qy 2312 TCTCAGGCTTCATGAGAAATTTCAATTTCACTCTGAAGATGAAATTAATGAACCAAA 2371
Dy 2288 TCTCAGGCTTCATGAGAAATTTCAATTTCACTCTGAAGATGAAATTAATGAACCAAA 2347
Qy 2372 AAAGCTG-----GAAGAGAGGAGGACTTGAATGTCTTACATTTTCAAGATCTTCTT 2425
Dy 2348 AGAGCTGGCAGAGAGAGAGAGGAGGAGATTTGAACGTCTGACGTTTCAAGACCTCTT 2407
Qy 2426 GCTTTGCATATCAAGTTGCCAAAGGAATTTCTGGAATTTAAAGTCGTGTGTACA 2485
Dy 2408 GCTTTGCATATCAAGTTGCCAAAGGAATTTCTGGAATTTAAAGTCGTGTGTACA 2467
Qy 2486 GAGACCTGGCCGCGCAGAACGCTGTTGTCACCCACGAGGAAAGTGGTGAAGATATGTGACT 2545
Dy 2468 GAGACCTGGCCGCGCAGAACGCTGTTGTTGTCACCCACGAGGAGTGGTGAAGATCTGTGACT 2527
Qy 2546 TTGGAATGGCTCGAGATATCATGATGATTTCCAACTATGTTGTCAGGGCAATGCCCTC 2605
Dy 2528 TTGACCTGGCCGCGCAGAACGCTGTTGAGCGACTCCAGCTAGCTGAGGGCAACGACGCG 2587
Qy 2606 TGCTGTAAATGATGGCCGCGCAGAACGCTGTTGAGGAGATCTACACCATTAAGAGTG 2665
Dy 2588 TGCCGCTGAAGTGCATGGCACCAGAGACTTTATTTGAAGGATCTACACATCAAGAGTG 2647
Qy 2666 ATGCTGGTCTATGGAATATTTACTGTGGGAAATCTTCTCAGCTGGTGTGAATCTTACC 2725
Dy 2648 ACCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
Qy 2726 CTGSCATTCGGTTGATGTAACCTCTACAACTGATTCAAANTGGATTTAAATGGATC 2785
Dy 2708 CTGSCATTCCTGTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAATGGAGC 2767
Qy 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGAATGCAATCTCTGGGCTTTTGACT 2845
Dy 2768 AGCCATTTATGCCACAGAGGATATACTTTGTAATGCAATCTCTGGGCTTTTGACT 2827
Qy 2846 CAAGGAAACGCCATCTTCCCTAATTTGACTTCGTTTTTAGGATGTACGTGGCAGATG 2905
Dy 2828 CAAGGAAACGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTACGTGGCAGAGG 2887
Qy 2906 CAGAGAGCGCATGTATCAGATGTCGATGCGCTGTTTCCGAATGTCCTCACACCTACC 2965
Dy 2888 CAGAAGAAG-----CATGTATCAGAAACATCCATCCATCTACC 2924
Qy 2966 AAAACGCGCACTTTTCAGCAGAGAGATGGATTTGGGCTACTCTCCGAGGCTCAGG 3025
Dy 2925 AAAACGCGCGCTTCAGCAGAGAGAGGCGCTCAGAGC--CCAGTCCGACAGCGCGCAGG 2983
Qy 3026 TCGAAGATTCGTAGAGGAAATTTAGTTTTTAAAGACTTTCATCTCCACCTATCCCTTAA 3085
Dy 2984 T-GAAGATTCACAGAGAAAGAGTTAGCGAGAGGCGCTTGGACCCCGC---CACCCTAG 3038
Qy 3086 CAGCTGTAGATTACCAAAACAGATTAATTTCACTACTAAAGAAAATCTATTATCAAC 3145
Dy 3039 CAGCTGTAGACCGCAGAGCAAGATTAGCTCGCTCT--GAGGAAGCGCCCTACACGG 3096
Qy 3146 TGCTGCTTACCAGACTTTTCTCTAGAGCGCTGCTGCTTACTCTTGTGTTTCAAGGGA 3205

Db 3097 CTTGCTTGGCTGGAGCTTTTCTCTAGATGCTGCTGCGCAATTACTC-----CAGAAGTGA 3149
QY 3206 CTTTGTAAATCAATCATCTCTGCACAAAGCAGGAGAGCTGATAATGAACCTTTATTG 3265
Db 3150 CTTCTATAAAATCAAAACCTCTCTCGCACAGCGGGAGGCAATAATGAGACTCTTTGG 3209
QY 3266 GAGCAATGATCGATCCAAAGCCCTTCTCAGCGCGCTTGAGTGAATTTGTGTACCTGAAG 3325
Db 3210 TGAGCCGCGCTACCTCGGGGCTTTCCAGAGCTTGAGGGGAAAGCCATGATCTGAAA 3269
QY 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTAGGAGAAGC 3378
Db 3270 TATAGTATATCTTGTAAATACGTGAACAAACCAACCCGCTTTTGTAGGGAAGC 3329
QY 3379 TAATATGATTTTTT--AAGCTATGTTTTAAATAATATGTAATTTTTTCAGCTATTAG 3436
Db 3330 TAAATATGATTTTTTAAATATCATGTTTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
QY 3437 TGATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAAAAAAAAAAAA 3496
Db 3390 TGATATATTTATGATGGAATAAACTTTCTACTGTAAAAAATAAAAAAAAAAAAAA 3449
QY 3497 AAAA 3500
Db 3450 AAAA 3453

RESULT 6
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-17
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-05-01
; PRIOR FILING DATE: 2000-05-01
; PRIOR FILING DATE: 2000-05-04
; PRIOR FILING DATE: 2000-05-22
; PRIOR FILING DATE: 2000-05-22
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7538

Query Match 18.1%; Score 633; DB 9; Length 645;

Best Local Similarity 99.8%; Pred. No. 1.3e-152;
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2059 CTGGGAAGCCAGAGAAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 2118
Db 1 CTGGGAAGCCAGAGAAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 60
QY 2119 TACTTGATTTTTGAATPACTGTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
Db 61 TACTTGATTTTTGAATPACTGTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
QY 2179 GAAAAATTTACAGAGACTTTGGACAGAGATTTTCAAGGAACACAAATTTTCAGTTTTTACCC 2238
Db 121 GAAAAATTTACAGAGACTTTGGACAGAGATTTTCAAGGAACACAAATTTTCAGTTTTTACCC 180
QY 2239 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGCTTCAAGAGAAGTTCAGATACACCCG 2298
Db 181 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGCTTCAAGAGAAGTTCAGATACACCCG 240
QY 2299 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAA 2358
Db 241 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAA 300
QY 2359 TATGAAACCAAAAAGGCTGGAAGAGAGAGACTTGAATGTGCTTACATTTGAAGAT 2418
Db 301 TATGAAACCAAAAAGGCTGGAAGAGAGAGACTTGAATGTGCTTACATTTGAAGAT 360
QY 2419 CTTCTTTTGGTTTGCATATCAATGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTG 2478
Db 361 CTTCTTTTGGTTTGCATATCAATGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTG 420
QY 2479 GTTCACAGAGACCTGGCCGCCAGAGAGCTGCTGTCAACCCACGGGAAAGTGGTGAAGATA 2538
Db 421 GTTCACAGAGACCTGGCCGCCAGAGAGCTGCTGTCAACCCACGGGAAAGTGGTGAAGATA 480
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACAGTATTCACACTATGTTGTCAGGGCAAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACAGTATGTTGTCAGGGCAAT 540
QY 2599 GCCCGTCTGCTCTAAATGGATGGCCCGCCGAAAGCCCTG-TTTGAAGGCATCTACACCAT 2657
Db 541 GCCCGTCTGCTCTAAATGGATGGCCCGCCGAAAGCCCTGTTTGAAGGCATCTACACCAT 600
QY 2658 TAAGAGTGTGCTGTGCTATATGATATTAATTAATTAATTAATTAATTAATTAATTAAT 2702
Db 601 TAAGAGTGTGCTGTGCTATATGATATTAATTAATTAATTAATTAATTAATTAATTAAT 645

RESULT 7
US-10-040-862-7538
; Sequence 7538, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-17
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28

```

; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7538

Query Match      18.1%; Score 633; DB 9; Length 645;
Best Local Similarity 99.8%; Pred. No. 1.3e-152;
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2059 CTGGGAAGCCACGAGAATATTGTAACCTGCTGGGGCGTGACACTGTGACGACCAATT 2118
Db 1 CTGGGAAGCCACGAGAATATTGTAACCTGCTGGGGCGTGACACTGTGACGACCAATT 60

QY 2119 TACTTGATTTTGAATACTGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGA 2178
Db 61 TACTTGATTTTGAATACTGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGA 120

QY 2179 GAAAAATTCACAGACTTGGACAGAGATTTCAAGGAACACAAATTTCAAGTTTACCCC 2238
Db 121 GAAAAATTCACAGACTTGGACAGAGATTTCAAGGAACACAAATTTCAAGTTTACCCC 180

QY 2239 ACTTTCCAATCAATCCAAATTCAGGCTTCCAGATGCTTCAAGGAAGTAAAGAT 2418
Db 181 ACTTTCCAATCAATCCAAATTCAGGCTTCCAGATGCTTCAAGGAAGTAAAGAT 360

QY 2419 CTTCTTTGCTTTGCATATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAGTCGTG 2478
Db 361 CTTCTTTGCTTTGCATATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAGTCGTG 420

QY 2479 GTTCACAGAGACTGGCGCCAGGACGCTGTTGCACCCACGGGAAAGTGTGAAGATA 2538
Db 421 GTTCACAGAGACTGGCGCCAGGACGCTGTTGCACCCACGGGAAAGTGTGAAGATA 480

QY 2539 TGTGACTTTGGATTCGCTGAGATATCATGACTGATTCGAATATCTTCTCAGGGCAAT 2598
Db 481 TGTGACTTTGGATTCGCTGAGATATCATGACTGATTCGAATATCTTCTCAGGGCAAT 540

QY 2599 GCCCTCTGCCTGTAATAATGGATGGCCGCCGAAAGCCTC-TTTGAAGGCATCTACACCAT 2657
Db 541 GCCCTCTGCCTGTAATAATGGATGGCCGCCGAAAGCCTCTTTTGAAGGCATCTACACCAT 600

QY 2658 TAAGAGTGATGCTCGGTGCATATGGAATATTACTGTGGGAATCTT 2702
Db 601 TAAGAGTGATGCTCGGTGCATATGGAATATTACTGTGGGAATCTT 645
```

```

RESULT 8
US-09-796-692-8085
; Sequence 8085, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8085

Query Match      17.4%; Score 608.4; DB 9; Length 610;
Best Local Similarity 99.8%; Pred. No. 2.8e-146;
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2059 CTGGGAAGCCACGAGAATATTGTAACCTGCTGGGGCGTGACACTGTGACGACCAATT 2118
Db 1 CTGGGAAGCCACGAGAATATTGTAACCTGCTGGGGCGTGACACTGTGACGACCAATT 60

QY 2119 TACTTGATTTTCAATACACTGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGA 2178
Db 61 TACTTGATTTTCAATACACTGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGA 120

QY 2179 GAAAAATTCACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTACCCC 2238
Db 121 GAAAAATTCACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTACCCC 180

QY 2239 ACTTTCCAATCAATCCAAATTCAGGCTTCCAGATGCTTCAAGGAAGTAAAGAT 2298
Db 181 ACTTTCCAATCAATCCAAATTCAGGCTTCCAGATGCTTCAAGGAAGTAAAGAT 240

QY 2299 GACTCGGATCAATCTCAGGCTTTCATGGAATTCATTTCACTCTGAAGATGAATGAA 2358
Db 241 GACTCGGATCAATCTCAGGCTTTCATGGAATTCATTTCACTCTGAAGATGAATGAA 300

QY 2359 TATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTGAATGCTTACATTTGAAGAT 2418
Db 301 TATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTGAATGCTTACATTTGAAGAT 360
```

```
QY 2419 CTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGAAATTTAAGTCGTGT 2478
Db 361 CTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGAAATTTAAGTCGTGT 420
QY 2479 GTTCACAGAGACCTGGCGCCAGGAACGTTGTCACCCACGGGAAAGTGTGAAGATA 2538
Db 421 GTTCACAGAGACCTGGCGCCAGGAACGTTGTCACCCACGGGAAAGTGTGAAGATA 480
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTCAGGGGCAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTCAGGGGCAT 540
QY 2599 GCCCGTCTGCCCTGTAATGGAATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTACACCAAT 2658
Db 541 GCCCGTCTGCCCTGTAATGGAATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTACACCAAT 600
QY 2659 AAGAGTGATG 2668
Db 601 AAGAGTGATG 610

RESULT 9
US-10-040-862-8085
; Sequence 8085, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8085

Query Match 17.4%; Score 608.4; DB 9; Length 610;
Best Local Similarity 99.8%; Pred. No. 2.8e-146;
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2059 CTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 2118
Db 1 CTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 60
QY 2119 TACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAAGTAAAGA 2178
Db 61 TACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAAAGTAAAGA 120
QY 2179 GAAAAATTCACAGACCTGGACAGAGATTTCAAGGACACAAATTCAGTTTTTACCCC 2238
Db 121 GAAAAATTCACAGACCTGGACAGAGATTTTCAAGGACACAAATTCAGTTTTTACCCC 180
QY 2239 ACTTTCCAATCATCCAAATCCAGCATGCCCTGGTTCAAGAGAAGTTCAGATACACCCG 2298
Db 181 ACTTTCCAATCATCCAAATTCAGCATGCCCTGGTTCAAGAGAAGTTCAGATACACCCG 240
QY 2299 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300
QY 2359 TATGAAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAATGCTTACATTTGAAGAT 2418
Db 301 TATGAAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAATGCTTACATTTGAAGAT 360
QY 2419 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478
Db 361 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 420
QY 2479 GTTCACAGAGACCTGGCGCCAGGAACGTTGTCACCCACGGGAAAGTGTGAAGATA 2538
Db 421 GTTCACAGAGACCTGGCGCCAGGAACGTTGTCACCCACGGGAAAGTGTGAAGATA 480
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTCAGGGGCAT 2598
Db 481 TGTGACTTTGGATTGGCTCGATATATCATGAGTGATTCACAACTATGTTGTCAGGGGCAT 540
QY 2599 GCCCGTCTGCCCTGTAATGGAATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTACACCAAT 2658
Db 541 GCCCGTCTGCCCTGTAATGGAATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTACACCAAT 600
QY 2659 AAGAGTGATG 2668
Db 601 AAGAGTGATG 610

RESULT 10
US-09-796-692-8132/c
; Sequence 8132, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
```

; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132

Query Match 17.3%; Score 606.4; DB 9; Length 608;
Best Local Similarity 99.8%; Pred. No. 9e-146;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2288 AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG 2347
Db |||||||
QY 608 AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG 549
Db |||||||
QY 2348 ATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAGAGGAGGACTTCAATTCGCTTA 2407
Db |||||||
QY 548 ATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAGAGGAGGACTTCAATTCGCTTA 489
Db |||||||
QY 2408 CATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAT 2467
Db |||||||
QY 488 CATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAT 429
Db |||||||
QY 2468 TTAAGTCGTGTGTACAGAGACCTGGCCGCCAGGACGCTTGTCTACCCACCGGAAAG 2527
Db |||||||
QY 428 TTAAGTCGTGTGTACAGAGACCTGGCCGCCAGGACGCTTGTCTACCCACCGGAAAG 369
Db |||||||
QY 2528 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG 2587
Db |||||||
QY 368 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG 309
Db |||||||
QY 2588 TCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCCGGAAAGCCTGTTTGAAGGCA 2647
Db |||||||
QY 308 TCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCCGGAAAGCCTGTTTGAAGGCA 249
Db |||||||
QY 2648 TCTACACCATTAAGAGTGATGCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAC 2707
Db |||||||
QY 248 TCTACACCATTAAGAGTGATGCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAC 189
Db |||||||
QY 2708 TTGGTGTGAATCCTTACCCTGGCATTCCGGTTGATGCTTAACCTCTACAACTGATTCAAA 2767
Db |||||||
QY 188 TTGGTGTGAATCCTTACCCTGGCATTCCGGTTGATGCTTAACCTCTACAACTGATTCAAA 129
Db |||||||
QY 2768 ATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTTAATGCAAT 2827
Db |||||||
QY 128 ATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTTAATGCAAT 69
Db |||||||
QY 2828 CCTGCTGGGCTTTTGACTCAAGAAACGCCCAATCTTCCCTAATTTGACTTCGTTTTTAG 2887
Db |||||||
QY 68 CCTGCTGGGCTTTTGACTCAAGAAACGCCCAATCTTCCCTAATTTGACTTCGTTTTTAG 9
Db |||||||
QY 2888 GATGTCAG 2895
Db |||||||
QY 8 GATGTCAG 1
Db |||||||

; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8132

Query Match 17.3%; Score 606.4; DB 9; Length 608;
Best Local Similarity 99.8%; Pred. No. 9e-146;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2288 AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG 2347
Db |||||||
QY 608 AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG 549
Db |||||||
QY 2348 ATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAGAGGAGGACTTGAATTCGCTTA 2407
Db |||||||
QY 548 ATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAGAGGAGGACTTGAATTCGCTTA 489
Db |||||||
QY 2408 CATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAT 2467
Db |||||||
QY 488 CATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAT 429
Db |||||||
QY 2468 TTAAGTCGTGTGTACAGAGACCTGGCCGCCAGGACGCTTGTCTACCCACCGGAAAG 2527
Db |||||||
QY 428 TTAAGTCGTGTGTACAGAGACCTGGCCGCCAGGACGCTTGTCTACCCACCGGAAAG 369
Db |||||||
QY 2528 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG 2587
Db |||||||
QY 368 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG 309
Db |||||||
QY 2588 TCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCCGGAAAGCCTGTTTGAAGGCA 2647
Db |||||||
QY 308 TCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCCGGAAAGCCTGTTTGAAGGCA 249
Db |||||||
QY 2648 TCTACACCATTAAGAGTGATGCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAC 2707
Db |||||||

RESULT 11
US-10-040-862-8132/c
; Sequence 8132, Application US/10040862
; Publication No. US2003007896A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander

```
Db      248  TCTACACCAATTAAGAGTGATGCTGTGCATATATGAATATTACTGTGGAAATCTTCTCAC 189
QY      2708  TTGGTGTGAATCCTTACCTGGCAATTCGGTGTGATGCTAACTTCTCAAACTGATTCAAA 2767
Db      188  TTGGTGTGAATCCTTACCTGGCAATTCGGTGTGATGCTAACTTCTCAAACTGATTCAAA 129
QY      2768  ATGGATTAAATAGGATCAGCCATTTTATGCTACAGAGAATAATACATTATAATGCAAT 2827
Db      128  ATGGATTAAATAGGATCAGCCATTTTATGCTACAGAGAATAATACATTATAATGCAAT 69
QY      2828  CTTGCTGGGCTTTGACTCAAGAAAGCGGCATCTTCCCTAAATTTGACTTCGTTTTAG 2887
Db      68  CTTGCTGGGCTTTGACTCAAGAAAGCGGCATCTTCCCTAAATTTGACTTCGTTTTAG 9
QY      2888  GATGTCAG 2895
Db      8  GATGTCAG 1

RESULT 12
US-09-796-692-8927
; Sequence 8927, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8927

Query Match      17.3%; Score 605; DB 9; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.1e-145;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2059  CTGGGAAGCCAGAGAAATATTGAACTGTGGGGGGGTGCACACTGTCAAGACCAAT 2118
Db      1  CTGGGAAGCCAGAGAAATATTGAACTGTGGGGGGGTGCACACTGTCAAGACCAAT 60
QY      2119  TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
```

```
Db      61  TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
QY      2179  GAAAAATTTCAAGAGACTTGGACAGAGATTTTCAAGSAAACACAATTTCAAGTTTACCCC 2238
Db      121  GAAAAATTTCAAGAGACTTGGACAGAGATTTTCAAGSAAACACAATTTCAAGTTTACCCC 180
QY      2239  ACTTTCCAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAAAGTTTCAAGATACACCCG 2298
Db      181  ACTTTCCAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAAAGTTTCAAGATACACCCG 240
QY      2299  GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCTCACTCTGAAGATGAAATTTGAA 2358
Db      241  GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCTCACTCTGAAGATGAAATTTGAA 300
QY      2359  TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGCTTACATTTGAAGAT 2418
Db      301  TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGCTTACATTTGAAGAT 360
QY      2419  CTTCTTTGCTTTGCAATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 2478
Db      361  CTTCTTTGCTTTGCAATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 420
QY      2479  GTTACAGAGACTGCGCCGCCAGGACGCTTGTCCACCCAGGGGAAAGTGTGAAGATA 2538
Db      421  GTTACAGAGACTGCGCCGCCAGGACGCTTGTCCACCCAGGGGAAAGTGTGAAGATA 480
QY      2539  TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCACAACTATGTTCTCAGGGGCAAT 2598
Db      481  TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCACAACTATGTTCTCAGGGGCAAT 540
QY      2599  GCCCGTCTGCTGTAAATGGAATGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAAT 2658
Db      541  GCCCGTCTGCTGTAAATGGAATGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAAT 600
QY      2659  AAGAG 2663
Db      601  AAGAG 605

RESULT 13
US-10-040-862-8927
; Sequence 8927, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
```

; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8927

Query Match 17.3%; Score 605; DB 9; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.1e-145;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2059 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATT 2118
Db 1 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATT 60
|||||
QY 2119 TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
Db 61 TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
|||||
QY 2179 GAAAAATTCACAGGACTTGCACAGAGATTTTCAAGGAACAATTTTCAGTTTACCC 2238
Db 121 GAAAAATTCACAGGACTTGCACAGAGATTTTCAAGGAACAATTTTCAGTTTACCC 180
|||||
QY 2239 ACTTTCACATCAATCAAAATCCAGCATGCTGTTCAAGAGAGTTTCATACACACCG 2298
Db 181 ACTTTCACATCAATCAAAATCCAGCATGCTGTTCAAGAGAGTTTCATACACACCG 240
|||||
QY 2299 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358
Db 241 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300
|||||
QY 2359 TATGAAACCAAAAAGCGTGGGAAGAGAGAGACTTTGAATGCTTACATTTGAAGAT 2418
Db 301 TATGAAACCAAAAAGCGTGGGAAGAGAGAGACTTTGAATGCTTACATTTGAAGAT 360
|||||
QY 2419 CTCTCTTCTGCTTGCATATCAAGTTGCCAAAGGATGAAATTTCTGGAAATTTAAGTCGT 2478
Db 361 CTCTCTTCTGCTTGCATATCAAGTTGCCAAAGGATGAAATTTCTGGAAATTTAAGTCGT 420
|||||
QY 2479 GTTCACAGAGACCTGGCGGCCAGGAAGCTGCTTCCACCCACGGGAAAGTGGTGAAGATA 2538
Db 421 GTTCACAGAGACCTGGCGGCCAGGAAGCTGCTTCCACCCACGGGAAAGTGGTGAAGATA 480
|||||
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTCAGGGGCAAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTCAGGGGCAAT 540
|||||
QY 2599 GCCCGTCTGCTGTAATGATGGATGGCCCCCGGAAGCCCTGTTGAAGGCATCTACACATT 2658
Db 541 GCCCGTCTGCTGTAATGATGGATGGCCCCCGGAAGCCCTGTTGAAGGCATCTACACATT 600
|||||
QY 2659 AAGAG 2663
Db 601 AAGAG 605
|||||

RESULT 14
US-09-796-692-8273
; Sequence 8273, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Aigier, Alexander
; APPLICANT: Aigier, Paul A.
; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8273

Query Match 17.1%; Score 598.4; DB 9; Length 601;
Best Local Similarity 99.7%; Pred. No. 1e-143; 2; Indels 0; Gaps 0;

Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2059 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATT 2118
Db 1 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATT 60
|||||
QY 2119 TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
Db 61 TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
|||||
QY 2179 GAAAAATTCACAGGACTTGGACAGAGATTTTCAAGGAACAATTTTCAGTTTACCC 2238
Db 121 GAAAAATTCACAGGACTTGGACAGAGATTTTCAAGGAACAATTTTCAGTTTACCC 180
|||||
QY 2239 ACTTTCACATCAATCCAAATTCACAGCATGCTGTTCAAGAGAGTTTCAGATACACCG 2298
Db 181 ACTTTCACATCAATCCAAATTCACAGCATGCTGTTCAAGAGAGTTTCAGATACACCG 240
|||||
QY 2299 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358
Db 241 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300
|||||
QY 2359 TATGAAACCAAAAAGCGTGGGAAGAGAGAGACTTTGAATGCTTACATTTGAAGAT 2418
Db 301 TATGAAACCAAAAAGCGTGGGAAGAGAGAGACTTTGAATGCTTACATTTGAAGAT 360
|||||
QY 2419 CTCTCTTCTGCTTGCATATCAAGTTGCCAAAGGATGAAATTTCTGGAATTTAAGTCGT 2478
Db 361 CTCTCTTCTGCTTGCATATCAAGTTGCCAAAGGATGAAATTTCTGGAATTTAAGTCGT 420
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:00:43 ; Search time 4597.52 Seconds

(without alignments)
12332.831 Million cell updates/sec

Title: US-09-919-408-3

Perfect score: 3501

Sequence: 1 CGAGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_eston:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_fod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	650.6	18.6	714	13	BI461248 603206574
2	509.6	14.6	621	13	BI360262 387099 MA
3	486.2	13.9	765	9	AI323253 mp88b02.y
4	440	12.6	604	12	BF523018 UI-R-C2p-
5	437.8	12.5	460	10	AV713950 AV713950
6	402.6	11.5	618	9	AA120050 mp88b02.r

7	400.4	11.4	455	13	BM147250
8	389.6	11.1	507	13	BM484050
9	340.4	9.7	3312	11	AK004947
10	325.6	9.3	445	10	BE651447
11	318.6	9.1	5174	11	BC026713
12	314.4	9.0	522	10	BB644407
13	305.4	8.7	326	13	BM144393
14	299.6	8.6	646	10	AW772610
15	283.4	8.1	659	10	AW057705
16	274.8	7.8	568	10	BE679272
17	249.2	7.1	515	14	BO556177
18	240.4	6.9	394	12	BG145808
19	236.6	6.8	660	10	BE047675
20	230.4	6.6	276	12	BF176247
21	228.4	6.5	739	12	BF159205
22	228	6.5	480	9	AI150354
23	223.8	6.4	666	13	BG934591
24	218.6	6.2	346	13	BM253635
25	218.6	6.2	346	13	BM253796
26	218.2	6.2	503	13	BI481351
27	216.4	6.2	832	13	BI694328
28	214	6.1	337	9	AI556371
29	213.6	6.1	794	13	BI111563
30	212.8	6.1	347	13	BM253661
31	212.2	6.1	407	10	AV596131
32	210.4	6.0	940	9	AI325368
33	209.6	6.0	747	9	AI323643
34	204.8	5.8	702	12	BF163671
35	204.8	5.8	848	13	BI107537
36	202.6	5.8	669	13	BJ012204
37	201.8	5.8	531	12	BF320739
38	201.6	5.8	666	9	AL047744
39	200.8	5.7	983	12	BF782576
40	197.8	5.6	201	10	AW059681
41	197.4	5.6	412	12	BE853746
42	196.6	5.6	461	9	AI092636
43	194.8	5.6	575	12	BF930052
44	194.4	5.6	696	14	BM951127
45	193.8	5.5	665	13	BG922846

ALIGNMENTS

RESULT 1	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
LOCUS	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
DEFINITION	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
ACCESSION	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
VERSION	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
KEYWORDS	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
SOURCE	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
ORGANISM	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
REFERENCE	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
AUTHORS	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
TITLE	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
JOURNAL	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001
COMMENT	BI461248	BI461248	603206574F1 NIH_MGC_97 Homo sapiens	714 bp	mRNA	linear	EST 21-AUG-2001

Human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 714)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11687 row: c column: 19
High quality sequence stop: 709.
Location/Qualifiers

LOCUS	BI360262	3870599	MARC 2P1G	Sus scrofa	CDNA 5'	621 bp	linear	EST 01-AUG-2000
DEFINITION	BI360262.1 GI:15056290							
ACCESSION	EST.							
VERSION	piq.							
KEYWORDS	Sus scrofa							
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.							
ORGANISM	1 (bases 1 to 621)							
REFERENCE	Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.							
AUTHORS	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine							
TITLE	Unpublished (2000)							
JOURNAL	Contact: Smith TPL							
COMMENT	USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@mail.marc.usda.gov							
FEATURES	Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.							
source	PCR Primers							
	FORWARD: AGGAACAGCATGATGACCAT							
	BACKWARD: GTTTCCTCAGTCAGCAGC							
	Plate: 139 row: H column: 2							
	Seq primer: ATTTAGCTGACACTATAG.							
	Location/Qualifiers							
	1..621							
	/organism="Sus scrofa"							
	/db_xref="taxon:9823"							
	/clone_lib="MARC 2P1G"							
	/tissue_type="pooled"							
	/lab_host="DH10B"							
	/note="vector: pcwv sport6; Site_1: NotI; Site_2: Salt; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."							
BASE COUNT	190 a 132 c 141 g 158 t							
ORIGIN	Query Match 14.6%; Score 509.6; DB 13; Length 621; Best Local Similarity 88.9%; Pred. No. 4.4e-76; Matches 551; Conservative 0; Mismatches 69; Indels 0; Gaps 0;							
Qy	1725	CGTTTTACCTCGCTAATTTGTCACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCT	1784					
Db	2	CGTTTTAACCATGCTAATTTGTCACAGTACAAAAGCGATTCGGTACGAAAGCCAGCT	61					
Qy	1785	ACAGATGGTACAGGTGACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGA	1844					
Db	62	GCAGATGGTCCAGGTGACGGGGTCCCTGGATAACGACTACTTCTACATCGACTTCAGAGA	121					
Qy	1845	ATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTTACT	1904					
Db	122	ATATGAATATGATCCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTTACT	181					
Qy	1905	AGGATCAGGTGCTTTTGGAAAGTATGATCAAGCGCAACGCTTATGGAATTACCAACAGG	1966					
Db	182	GGGATCCGGCGCTTTTGGAAAGTATGATCAAGCGCAACGCTTATGGAATTACCAACGCT	241					
Qy	1965	AGTCTCAATCCAGGTTCGGTCAAAATGCTCAAAAGAAAGCAACGACTCTCGAAAGAGA	2024					
Db	242	AGTGTCAATCCAGGTTCGGTCAAAATGCTCAAAAGAAAGCAACGACTCTCGAAAGAGA	301					
Qy	2025	GGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAA	2084					
Db	302	GGCTCTCATGTCTGAACCTCAAAATGATGACCCACCTGGGAGCCAGCAGAAATATAGTAA	361					
Qy	2085	CCTGCTGGGGCGTGACACACTGTCTCAGGACCAGTATTTTGAATCTGTTGCTA	2144					

Qy 2085 CCTGCTGGGGGGCTGCACACTGTCAGGACCAATTACTGTGATTTTGAATACTGTGTGCTA 2144

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 604)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
9704477

JOURNAL
MEDLINE
COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1793119
Seq primer: M13 Forward.

FEATURES
Source

1. 604
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-Ig-f-10-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2p library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"
BASE COUNT 166 a 132 c 167 g 139 t
ORIGIN

Query Match 12.6%; Score 440; DB 12; Length 604;
Best Local Similarity 85.4%; Pred. No. 2.2e-64;
Matches 504; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

QY 2016 TGAAGAGAGGAGGACTCATGTCAGAACTCAAGATGATGACCCAGCTGGAGAGCCAGAGAA 2075
DB 9 TGAAGAGAGGCGCTCATGCTGAGCTCAAAATGATGACCCAGCTGGGAGCCAGATGACAA 68

QY 2076 TATTGTGAACCTCTGGGGCGTGCACACTGTGAGGACCAATTTACTTGATTTTGAATA 2135
DB 69 CATCGTGAACCTCTGGGGCGATGCACACTGTGAGGCGCAGTGTGATTTTGAATA 128

QY 2136 CTGCTCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAGAAAAATTTACAGGAC 2195
DB 129 TTGTTGCCATGGTGACCTGCTCAACTACCTAAGAGCAAAAGAGAGAAAGTTTACAGGAC 188

QY 2196 TTGGACAGAGATTTTCAAGAGAACACAAATTTTACGTTTTTACCCCACTTTTCAATCACATCC 2255
DB 189 GTGGACAGAGATTTTAAAGGAACATAATTTTACGTTTTTACCCCACTTTTCCAGTCACATTC 248

QY 2256 AAATTCAGCATGCTGCTGTTCAAGAGAGAGTTCAGATACACCCGGAGTCCGATCAAAATCTTC 2315
DB 249 AAATCCAGTATGCGCGGTTTCAGAGAGAGTTCAGATATACCCCGCCGTCAGTCAAGTCTC 308

QY 2316 AGGGCTTCATGGGAATTCATTTCACTCTCAAGATGAATTAATATGAATAACCAAAAG 2375
DB 309 AGGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAGTATGAATAACCAAGAG 368

QY 2376 GCTGGAAGAGAGGAG- - - - -GACTTGAATGTGCTTACATTTGAAGATCTTCTTTGCTT 2429
DB 369 GCTGGAAGAGAGGAGGAGAGATTTGAACGTGCTGACGTTTGAAGACCTCTCTTTGTT 428

QY 2430 TCATATCAAGTTGCCAAAGAGATTTCTTGGAAATTTAAGCTGTGTGTTCCACAGAGA 2489
DB 429 TCGTATCAAGTGGCCAGGATGAGTTCTCTGGAGTTCAAGTCTGTGTTCCACAGAGA 488

QY 2490 CTTGGCCGCCAGGAGAGCTGCTGTCAACCCAGGAGAGTGTGAAGATATGTCATTTGG 2549
DB 489 CTTGGCAGCCAGGAGAGCTACTGTGTCAACCCAGGAGAGTGTGAAGATCTGTGACTTTGG 548

QY 2550 ATTGGCTCGAGATATCATGATGATTCACAACTATGTTGTCAGGGGCAATG 2599
DB 549 ACTGGCCCGAGACATCTGAGTGACTCCAGCTACGTTCTTCAGGGGCAACG 598

RESULT 5
AV713950 460 bp mRNA linear EST 11-OCT-2000
LOCUS AV713950 DCB Homo sapiens cDNA clone DCBADE05 5', mRNA sequence.
DEFINITION AV713950
ACCESSION AV713950
VERSION AV713950.1 GI:10795467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. 460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DCBADE05"
/clone_lib="DCB"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/note="Vector: pTriplex2; Site_1: sf1RA; Site_2: sf1fB"
BASE COUNT 152 a 95 c 96 g 117 t
ORIGIN

Query Match 12.5%; Score 437.8; DB 10; Length 460;
Best Local Similarity 99.5%; Pred. No. 5.8e-64;
Matches 439; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1930 ATGAACCAACAGCTTATGGAATTAGCAAAACAGAGTCTCAATCCAGGTTCGCTCAAA 1989
DB 1 ATGAACCAACAGCTTATGGAATTAGCAAAACAGAGTCTCAATCCAGGTTCGCTCAAA 60

QY 1990 ATGCTGAAGAAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGACTCAAGATG 2049

Db 61 ATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAAGACTCAAGATG 120
 QY 2050 ATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGTGACACTGTCA 2109
 Db 121 ATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGTGACACTGTCA 180
 QY 2110 GGACCAATTTACTTGATTTTGAATCTGTTGCTATGTGTATGCTTCTCAACTATCTAAGA 2169
 Db 181 GGACCAATTTACTTGATTTTGAATCTGTTGCTATGTGTATGCTTCTCAACTATCTAAGA 240
 QY 2170 AGTAAAGAGAAATTTTCACAGACTTGGACAGAGATTTTCAAGGAACAATTTTCAGT 2229
 Db 241 AGTAAAGAGAAATTTTCACAGACTTGGACAGAGATTTTCAAGGAACAATTTTCAGT 300
 QY 2230 TTTTACCCCACTTTCCCAATCACATCCAAATTCACAGATGCTGCTCAAGAGAAGTTTCAG 2289
 Db 301 TTTTACCCCACTTTCCCAATCACATCCAAATTCACAGATGCTGCTCAAGAGAAGTTTCAG 360
 QY 2290 ATACACCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGAT 2349
 Db 361 ATACACCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGAT 420
 QY 2350 GAAATTGAATGAAACCCAA 2370
 Db 421 GAAATTGAATGAAACCCAA 441

RESULT 6

AA120050 618 bp mRNA linear EST 17-FEB-1997
 LOCUS mp88b02.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:576267
 DEFINITION 5', similar to gb:M64689 Mouse flk-2 mRNA, complete cds (MOUSE),,,
 mRNA sequence.

ACCESSION

VERSION AA120050.1 GI:1677559

KEYWORDS

SOURCE EST.

ORGANISM

house mouse.
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 618)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE

The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 335.

Location/Qualifiers

1..618
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:576267"
 /clone_lib="Soares_thymus_2NbMT"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not 1 - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGAGCGCCGGTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not 1 and cloned into the Not 1
 and Eco RI sites of the modified pT7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 158 a 142 c 158 g 160 t
 ORIGIN

Query Match 11.5%; Score 402.6; DB 9; Length 618;
 Best Local Similarity 82.4%; Pred. No. 4.3e-58;
 Matches 462; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 QY 1582..GGGTTCCTGGTCAAGTGTGTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCCTT 1641
 Db 1 GGGCTTCTGGTCAAAATGCTGTGCGTACAATTTATGGGCACGCTCTTGGCAAAACCATCTTT 60
 QY 1642 TTAACCTCTCCAGGCCCCCTCCCTTTCATCCAAAGACACATCTCATCTATGCAACAAT 1701
 Db 61 TTAACCTCACCAGGCCCCCTCCCTTTCATCCAAAGACACATCTCATCTATGCGACCAT 120
 QY 1702 GGTGTTTCTCTCTCTTCATCTGCTTTTAAACCTGCTAAATTTGTCAAGTACAAAAAG 1761
 Db 121 GGGCTCTCTCCCTCTCATTTGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAAG 180
 QY 1762 CAATTTAGTATGAAAGCCAGCTACAGATGTTACAGGTGACCGGTCTCCCTCAGATAATGAG 1821
 Db 181 CAATTTAGTACGAGAGTCAAGTGCAGATGATCCAGGTGACCTGGCCCCCTGGATAACGAG 240
 QY 1822 TACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTCAAAATGGAGGTTTCCAAGAGAA 1881
 Db 241 TACTTCTACGTTGATTTTCAGGAGCTATGAATATGATCTCAAAATGGAGGTTTCCAAGAGAG 300
 QY 1882 AATTTAGAGTTTGGGAAGGTACTAGATCAGGTGCTTTTGGAAAAGTGTGATGACGCAACA 1941
 Db 301 AACTTAGAGTTTGGGAAGGTCTCGGGTCTGGCGCTTTTCGGAGGTGATGACGCCACG 360
 QY 1942 GCTTATGAATTAGCAAAACAGAGTCTCAATCCAGTTTGGCGTCAAAATGCTGAAAGAA 2001
 Db 361 GCCTATGCAATTAGTAAACGGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAG 420
 QY 2002 AAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGACCCAGCTG 2061
 Db 421 AAAGCTGACAGCTGTGAAAAGAGAGCTCTCATGTGCGAGCTCACAATGATGATGACCCACCTG 480
 QY 2062 GGAAGCCACGAGAAATTTGTAACCTGCTGGGGCGTGCACACTGTCAGGACCAATTTAC 2121
 Db 481 GGACACCATGACACATCGTAATCTGCTGGGGCATGCACACTGTCAGGCGCAGTGTAC 540
 QY 2122 TTGATTTTGAATCTGTTGC 2142
 Db 541 CTGATTTTGAATGTTGC 561

RESULT 7

BM147250

LOCUS

DEFINITION

sequence.

BM147250

BM147250.1 GI:17166911

EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 455)

REFERENCE

AUTHORS

Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,

Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.

BM147250 455 bp mRNA linear EST 30-NOV-2001
 TCAAP1010678 Pediatric acute myelogenous leukemia cell (FAB M1)
 Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1067, mRNA

TITLE Pediatric Leukemia' cDNA Sequencing Project (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Seq primer: M13 primer.

FEATURES
source Location/Qualifiers
1..455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAAP1067"
M1 Baylor-HGSC project=leukemia cell (FAB
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTCGAGGCCCGCAGGAG(T)VN
3'; V=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dc primer
[5'AGAGCTCGGTACCGCGCCGAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagakawa S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT	95 a	118 c	129 g	112 t	1 others
ORIGIN					
Query Match	11.4%	Score 400.4;	DB 13;	Length 455;	
Best Local Similarity	95.8%;	Pred. No. 1.1e-57;			
Matches 432;	Conservative 0;	Mismatches 17;	Indels 2;	Gaps 2;	
QY	1	CGAGCGGCATCGAGGCGTCGGCGCGGC-CCTGGGGGACCCC GGCTCCGGAGGCCAT	59		
Db	6	CGAGCGGCATCGAGGCGTCGGCGCGGCCTGGGACCCCGGCTCCGGAGGCCAT	65		
QY	60	GCGGCGGTTGGCGGCGACGGGGCACCGTGCCGCTGCTGCTGTTTTTCTGCAATGAT	i19		
Db	66	GCGGCGGTTGGCGGCGACGGGCGCAGCTGCCGCTGCTGCTGTTTTTCTGCAATGAT	125		
QY	120	ATTGGGACTATTACAATCAAGATGCGCTGATCAAGTGCTTTTAATCAATCATAA	179		
Db	126	ATTGGGACTATTACAATCAAGATGCGCTGATCAAGTGCTTTTAATCAATCATAC	185		
QY	180	GAACAATGATTATCACTGGGGAAGTCATCATATCCCCATGGTATCAGAATCCCCGGA	239		
Db	186	GAACAATGATTATCACTGGGGAAGTCATCATATCCCCATGGTATCAGAATCCCCGGA	245		
QY	240	AGACCTCGGGTGTGCGTTGAGACCCCGAGAGCTCAGGACATGTACGAAGCTGCGGTGT	299		
Db	246	AGACCTCGGGTGTGCG-TGAGACCCCGAGAGCTCAGGACATGTACGAAGCTGCGGTGT	304		
QY	300	GGAAGTGGATGTATCTGCTCCATCACACTGCAAGTCTGCTCGATGCCCGGAGAACAT	359		
Db	305	GGAAGTGGATGTATCTGCTCCATCACACTGCAAGTCTGCTCGATGCCCGGAGAACAT	364		
QY	360	TTCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGTATTACA	419		
Db	365	TTCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGTATTACA	424		
QY	420	AAACAGAGGAGTGTGTTTCCATGGTCATT TTG	450		

[illegible]

QY 1493 CAGAAGGAGCTGGAATAGAAAGGCTAACAGAAAAGTGTGACACAGTGGGTGTCGAGCA 1552
 DB 300 CGGAAGGCATCTGGAATAAAGGCGCAACAGGAAGATATTGACACAGTGGATTCCAGCA 359
 QY 1553 GTACTCTAAACATGAGTGAAGCCATAAAAGGGTTCTCTGGTCAAGTCTGTGCATACAATT 1612
 DB 360 GCATCTGAACATGACGAGGCTGTCAAGGGTTCTGTGGAAGTCTGTGCGCTACAATT 419
 QY 1613 CCCTTGGACATCTTCTGTGACGATCCCTTTTAAACTCTCAGGCCCTTCCCTTTCATCC 1672
 DB 420 CCCTGGGACGCTTCTGTGAACGATCCCTCTTAAACTCACAGGCCCTTCCCTTTCATCC 479
 QY 1673 AGACAACATCTCATCTTATGCAACAAT 1700
 DB 480 AGACAACATCTCATCTTATGCAACAAT 507

RESULT 9
 AK004947
 LOCUS
 DEFINITION
 Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300008N20:colony stimulating factor 1 receptor, full insert sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library clone:1300008N20.
 ORGANISM
 Carninci,P. and Hayashizaki,Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 REFERENCE
 2
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159
 REFERENCE
 3
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 REFERENCE
 4
 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,N., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,

Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsaki,S. and Hayashizaki,Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 MEDLINE
 21085660
 PUBMED
 11217851
 REFERENCE
 5
 (bases 1 to 3312)
 Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGCGCGCGCACTCGAGTTTTTTTTTTTNN 3', cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGAGGATCCAGAGCTCAATTATTAACCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: 3' end: XhoI. Host: SOLR.
 FEATURES
 source
 Location/Qualifiers
 1..3312
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:1300008N20"
 /db_xref="MGD:MGI:1904396"
 /db_xref="taxon:10090"
 /clone="1300008N20"
 /sex="male"
 /tissue_type="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1..3312
 /gene="Csflr"
 93..3026
 /gene="Csflr"
 /note="colony stimulating factor 1 receptor data source:MGD, source key:MGI:1339758, evidence:ISS putative"
 /codon_start=1
 /protein_id="BAB23691.1"
 /db_xref="GI:12836518"
 /db_xref="MGD:MGI:1339758"
 /translation="MELGPPVLVLLATVWHGQAPVIEPSGPELVVEPGETVTLRCVSKNSVMDGPISPYWTLDPPSPGSLTTRNATKCTYCTELEDPMAGSTHLYVK DPAHSWNLLAQEVTVVEGQAVLPCLITDPAKDSVSLMREGGRQVLRKTVFFSPWR

cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996).

BASE COUNT	96 a	123 c	111 g	115 t
ORIGIN				

Query Match 9.3%; Score 325.6; DB 10; Length 445;
Best Local Similarity 83.3%; Pred. No. 4.3e-45;
Matches 370; Conservative 0; Mismatches 74; Indels 0;

637	QY	GTGGAATGGGTGCTTTGCGAATTCACAGGGGGAAGCTGTAAAGAAGAAAAGTCCAGCTGTT	696
444	Db	GTGGAGTGGGTGCTCTCGAGTCCACAGGGAAGCTGTAAAGAAGAAAGGCCCTGCTGTT	385
697	QY	GTTTAAAAGAGGAGNAAGTGCCTTCATGAATATTTCGGACGGACATAAAGTGCTGTGCC	756
384	Db	GTCAAGAAGAGGAGAAAGGTACTTCATGAGTGTTCGGAACAGACATCAGATGCTGTGCT	325
757	QY	AGAAATGAATGGCGAGGGAATGACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCT	816
324	Db	AGAAATGCACGTGGCGCGAATGCACCAAGCTGTTTACCATAGATCTAAACCCAGGCTCCT	265
817	QY	CAGACACATGGCCACAAATTATTTCCTTAAAGTAGGGGAACCTTATGGATAAGGTGCAAA	876
264	Db	CAGAGCACATGCCCCAGTTTATTCCTGAAAGTGGGGGAACCTTGTGGATCAGCTGTAAG	205
877	QY	GCCTGTTTCATGTGAACATGGATTGCGGCTCACCTGGGAAATTAGAAAAACAAGCACTCGAG	936
204	Db	GCCATCCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCCTGGAG	145
937	QY	GAGGCAACTACTTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGATTCTG	996
144	Db	GAGGGCAGCTACTTTGAGATCAGTACCTACTCCACAAACAGGACCATGATTCGGATTCTC	85
997	QY	TTTGTCTTTTCTATCATCAGTGGGAAGAAACGACACCGGATACTACACTTGTTCCTCTCA	1056
84	Db	TTGCGCTTTGTGCTCCCGTGGGAAGAACACACCCGGATATTACACTTGTCTTCTCTCA	25
1057	QY	AAGCATCCCAAGTCAATCAGCTTTG	1080
24	Db	AAGCACCCAGCAGCATCAGCGTTG	1

RESULT 11

BC026713

LOCUS BC026713

DEFINITION	Mus muscu
ACCESSION	PC026713

ACCESSION	BC026713
VERSION	BC026713

KEYWORDS HTC.

SOURCE house mov

•ORGANISM Mus muscu

Eukaryotae
Mammalia

REFERENCE

AUTHORS

TITLE
JOURNAL

REMARK
COMMENT

Direct submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: ang@bcm.tmc.edu
Gunsaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 45 Row: b Column: 17
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
 This clone has the following problem: no 5' EST match

FEATURES source

```

1356 a 1232 c 1246 g 1340 t
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N-3"
/clone="IMAGE:5008623"
/tissue_type="Mammary tumor
old mouse. Taken by biopsy.
/clone_lib="NCI-CCAP_Mam2"
/lab_host="DH10B"
/notes="pCMV-Sport6"
BASE COUNT

```

BASE COUNT
ORIGIN

Query Match 9.18; Score 318.6; DB 11; Length 5174;

Best Local Similarity	57.1%; Pred. No. 2.3e-44;				
Matches 601; Conservative	0; Mismatches 449; Indels	3; Gaps	1;		
QY	1816 AATGAGTACTCTACGTTTGATTTCAGAGAATAATGAATATGATCTCAAAATGGGAGTTCCCA	18175			
Db	1749 AACAAITATGTTTTACATAGACCAGCGRACCTTCTTATGATCACAAATGGGAGTTCCC	1808			
QY	1876 AGAGAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTTGGAAAAAGTATGAAC	19355			
Db	1809 AGAACAAGCGTGAGTTTTTGGAAAAGACATTTGGGAGCTGGTGCCTTTCGGGAAGGTCGTTGAG	1868			
QY	1936 GCNAACAGCTTATGGAAITTAGCAAACACAGAGTCTCAATCCAGGTTGCCGTCAAAATGCTG	1995			
Db	1869 GCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCATGACAGTTGCCGTGAAGATGCTC	1928			
QY	1996 AAAGAAAAAGCACAGCTCTGAAAAGAGGGCACTCATGTGCAGAACTCAAGATGATGACC	2055			
Db	1929 AAACCAAGTGCCCATTTTAAACAAGAGAGAGGCCCTAATGTGCGAACTGAAGGTCTCGAGC	1988			
QY	2056 CAGCTGGGAGGCCACAGAGAATATTGTGAACCTGCTGGGGCGGTGCACACTGTGCAGGACCA	2115			
Db	1989 TACCTGGGCAATCACATGAATATTGTGAACCTGCTTGGCGCATGCACGGTGGAGGGCCCC	2048			
QY	2116 ATTTACTTGATTTTGAATACTCTTGCTATGGTGATCTTCTCAACTATCTTAAGAAAGTAAA	2175			
Db	2049 ACCCTGGTCATTACAGAAATATTGTTGCTATGGTGATCTTTTGAATTTTTTGAAGAAGAAG	2108			
QY	2176 AGAGAAAATTTACAGGACTTGGACAGAGATTTTTCAAGAACACAAATTTTCATGTTTTTAC	2235			
Db	2109 CGTGACTCGTTTTATTTCTTCAAAGCAAGAAGAGCAGGCGAGCGGCACCTTTTATAAGAAC	2168			

Db 127 CAACCAAGACCTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCAT 186

Oy 194 CAGTGGGAAGTCAATCATATCATATCCATGATATCCAGTATCCCGAAGACCTCGGGTGTG 253

Db 187 CAGCGGGAAGCCATCATGTCGTAACGAATGTCGAGGATCCCGAGACCTCAGTGTG 246

Oy 254 CGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCGCGTGTGGAAGTGGATGAT 313

Db 247 CCGCGAGCGCGCAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGT 306

Oy 314 CTGCTTCATACACTGCAAGTGTGTCGATGCCCGGATGCCCGAGGAACTTTCCTGTCTCTGGG 373

Db 307 CTGGTCCATACACCTGCAAGTGTGAGCTGCGCCACCCAGGGGAGCTTTCCTGCTCTGGG 366

Oy 374 TCTTTAAGCAGCTCCCTGAAATGCCAGCCACACATTTTGATTTTCAAAACAGAGAGTGTG 433

Db 367 TCTTTAAGCAGCTCCCTGCGCTGCCAGCGCACTTTGATTTTCAAAACAGAGAGATCG 426

Oy 434 TTTCCATGGTCATTTTGAATATGACAGAACCCCAAGCTGGAGATACCTATTTTATTC 493

Db 427 TTTCCATGGCTATCTTGAAGCTGTACAGAGACCCAGGAGGAGATACCTACTCCATATTC 486

Oy 494 AGAGTGAAGCTACCAATTTACACATATTTGTTTACAG 529

Db 487 AGACGGAAGCGGCC-ACTACACAGTACTGTTTACAG 521

RESULT 13 BM144393/c

LOCUS BM144393 326 bp mRNA linear EST 30-NOV-2001
DEFINITION TCAAP1D10678 Pediatric acute myelogenous leukemia cell (FAB M1)
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1067, mRNA
sequence.

ACCESSION

VERSION BM144393.1 GI:17161450
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Wei,X., Tsang,Y.T.M., Mel,G., Ku,J.M., Ali-Osman,F.R. Jr.,

Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
Unpublished (2001)

JOURNAL

COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org

Seq primer: M13 primer.

Location/Qualifiers

1..326

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="TCAAP1067"

/clone_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project-TCAA"

/sex="male"

/tissue_type="leukopheresis"

/cell_type="myeloid cell"

/dev_stage="pediatric 6 years"

/lab_host="DH10B"

/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCGAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGAGCTCGGATCCGCGCGCAATAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 111 a 56 c 45 g 114 t
ORIGIN

Query Match 8.7%; Score 305.4; DB 13; Length 326;
Best Local Similarity 97.9%; Pred. No. 1.2e-41;
Matches 320; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Oy 3163 TTTCTCTAGAAGCCGCTCTGCTTTACTCTCTGTTTCAAGGAGCATTTCGTAATAAATCAAT 3222

Db 326 TTTCTCTACAAGCTGCTGCGTTTACTCTCTGTTTACAAGA-ACCTTTGTAAATCAAT 268

Oy 3223 CATCTGTCAACAAGCAGGAGGAGCTGATAATGAACCTTATTGGAGCATTGATCTGCATC 3282

Db 267 CATCTGTCAACAAGCAGGAGGAGCTGATAATGAACCTTATTGGAGCATTGATCTGCATC 208

Oy 3283 CAAGGCCTTCTCAGGCCGCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTGTGA 3342

Db 207 CAAGGCCTTCTCAGGCCGCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTGTGA 148

Oy 3343 AATACATAAACAAGCATTTCGTAAGGAGAGCTAATATGATTTTAACTATGT 3402

Db 147 AATACATAAACAAGCATTTCGTAAGGAGAGCTAATATGATTTTAACTATGT 88

Oy 3403 TTTAAATAATATGTAAATTTTTCAGCTATTAGTGATATATTTATGGTGGGAATAA 3462

Db 87 TTTAAATAATATGTAAATTTTTCAGCTATTAGTGATATATTTATGGTGGGAATAA 28

Oy 3463 ATTTCTACTACAGAAAAA 3489

Db 27 ATTTCTACTACAGAAAAA 1

RESULT 14
AW772610/c

LOCUS AW772610 646 bp mRNA linear EST 04-MAY-2000
DEFINITION hn77b03.x1 NCI-CGAP Kid11 Homo sapiens cDNA clone IMAGE:3033869 3'

Similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);, mRNA
sequence.

ACCESSION AW772610 GI:7704608
VERSION AW772610.1
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 646)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov
Seq primer: -400P from Gibco

High quality sequence stop: 462.
Location/Qualifiers

1..646

/organism="Homo sapiens"

/db_xref="taxon:9606"

Insert Length: 1584 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 431.

FEATURES
source
1. .659
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2542486"
/clone_lib="NCI_CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

BASE COUNT 204 a 131 c 119 g 205 t

ORIGIN
Query Match 8.1%; Score 283.4; DB 10; Length 659;
Best Local Similarity 98.0%; Pred. No. 4.5e-38;
Matches 287; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3185 TTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTGTCACAGGAGG 3244
|||||
Db 639 TTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTGTCACAGGAGG 580
|||||
QY 3245 AGCTGATGAACTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTT 3304
|||||
Db 579 AGCTGATGAACTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTT 520
|||||
QY 3305 GAGTGAATCTGTACCTGAAGTACAGTATTTCTGTAAATACATAAAACAAAGCATTT 3364
|||||
Db 519 GAGTGAATCTGTACCTGAAGTACAGTATTTCTGTAAATACATAAAACAAAGCATTT 460
|||||
QY 3365 TGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAATTT 3424
|||||
Db 459 TGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAATTT 400
|||||
QY 3425 TCAGCTATTTAGTCATATATTTTATGGTGGGTAATAAAATTTCTACTACAGAA 3477
|||||
Db 399 TCAGCTATTTAGTCATATATTTTATGGTGGGTAATAAAATTTCTACTACAGAA 347

Search completed: May 25, 2003, 11:51:19
Job time : 4599.52 secs

/clone="IMAGE:3033869"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 13223911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 201 a 130 c 116 g 199 t

ORIGIN
Query Match 8.6%; Score 299.6; DB 10; Length 646;
Best Local Similarity 98.7%; Pred. No. 8.7e-41;
Matches 302; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3172 AAGCGCTGCGTTTACTCTGTTTCAAGGAGCTTTGTAAATCAATCATCTGTC 3231
|||||
Db 646 AAGCTGCTGCGTTTACTCTGTTTCAAGGAGCTTTGTAAATCAATCATCTGTC 587
|||||
QY 3232 ACAAGCAGGAGGAGCTGATGAACTTATTGGAGCATTTGATCTGCATCCAGGCGCTT 3291
|||||
Db 586 ACAAGCAGGAGGAGCTGATGAACTTATTGGAGCATTTGATCTGCATCCAGGCGCTT 527
|||||
QY 3292 CTCAGCGCGCTTGAGTGAATCTGTACCTGAAGTACAGTATATTTCTGTAAATACATAA 3351
|||||
Db 526 CTCAGCGCTGCTGAGTGAATCTGTACCTGAAGTACAGTATATTTCTGTAAATACATAA 467
|||||
QY 3352 AACAAAGCATTTTGTAAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATA 3411
|||||
Db 466 AACAAAGCATTTTGTAAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATA 407
|||||
QY 3412 ATATGTAAATTTTTCAGCTATTTAGTCATATTTTATGGTGGGTAATAAAATTTCTACT 3471
|||||
Db 406 ATATGTAAATTTTTCAGCTATTTAGTCATATTTTATGGTGGGTAATAAAATTTCTACT 347
|||||
QY 3472 ACAGAA 3477
|||||
Db 346 ACAGAA 341

RESULT 15
AW057705/c
LOCUS
DEFINITION
659 bp mRNA linear EST 20-OCT-2000
wx02c12.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2542486 3',
similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AW057705
AW057705.1 GI:5933344
EST.
Homo sapiens
Organism
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 659)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: May 27, 2003, 14:18:14 ; Search time 61.0307 seconds
(without alignments)
2168.053 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTYPLLVFSAM PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Genes2_101002.*
1: /SID52/gcgdata/geneseq/geneq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	16	AA67816
2	5274	100.0	993	16	AA67536
3	5274	100.0	993	17	AA97419
4	5274	100.0	993	18	AA19873
5	5274	100.0	993	20	AA198617
6	5271	99.9	993	14	AA37503
7	5266	99.8	993	14	AA44995
8	5266	99.8	993	16	AA75961
9	5265	99.8	993	16	AA81869
10	5265	99.8	1167	14	AA31376

11	5159	97.8	983	19	AAW63588	Human receptor typ
12	5157.5	97.8	986	19	AAW63587	Human receptor typ
13	5157.5	97.8	986	19	AAW63589	Human receptor typ
14	5153.5	97.7	994	19	AAW63586	Human receptor typ
15	4533.5	86.0	1000	16	AAW81868	Flk2/flt3 tyrosine
16	4429.5	84.0	992	14	AA37502	Murine flk-2. Mus
17	4429.5	84.0	992	16	AA67815	Flk2 receptor prot
18	4429.5	84.0	992	16	AA67535	Mouse flk-2. Mus
19	4429.5	84.0	992	17	AA97418	Human foetal liver
20	4429.5	84.0	992	18	AAW19874	Murine flk-2 recep
21	4429.5	84.0	992	20	AAW08616	Murine flk-2 prote
22	4421.5	83.8	992	14	AA44994	Murine flk-2 recep
23	4421.5	83.8	992	14	AA31375	Murine flk-2. Mus
24	4404.5	83.5	992	13	AA328038	Murine flk-2. Mus
25	3323	63.0	665	19	AAW63585	Human receptor typ
26	2747	52.1	749	19	AAW78002	Protein pMON32390.
27	1640	31.1	481	15	AAW47579	Soluble Flk-2. Mu
28	1286	24.4	977	21	AA51322	Bovine c-Kit bk-1
29	1266	24.0	975	22	AA607144	Murine Kit/stem ce
30	1266	24.0	975	22	AA607148	Mutant murine Kit/
31	1251	23.7	976	22	AA607145	Human Kit/stem cel
32	1251	23.7	976	22	AA607149	Mutant human Kit/s
33	1251	23.7	976	22	AAU00375	Human stem cell gr
34	1219.5	23.1	1055	22	ABG15479	Novel human diagno
35	1218.5	23.1	972	23	AAU11936	Colony stimulating
36	1217.5	23.1	972	23	AAU11934	Colony stimulating
37	1217.5	23.1	972	23	AAU11940	Colony stimulating
38	1216.5	23.1	972	23	AAU79039	Human macrophage c
39	1216.5	23.1	972	23	AAU11941	Colony stimulating
40	1216.5	23.1	972	23	AAU73585	Colony stimulating
41	1215.5	23.0	972	23	AAU11935	Colony stimulating
42	1215.5	23.0	972	23	AAU11937	Colony stimulating
43	1211.5	23.0	972	23	AAU11938	Colony stimulating
44	1207.5	22.9	972	23	AAU11939	Colony stimulating
45	1189.5	22.6	2129	22	ABG15478	Novel human diagno

ALIGNMENTS

RESULT 1
AA67816
ID AAR67816 standard; Protein; 993 AA.
XX AAR67816;
XX AAR67816;
DT 18-AUG-1995 (first entry)
XX Flk2 receptor protein-tyrosine-kinase.
DE Flk2 receptor protein-tyrosine-kinase.
XX Human Flk2; receptor protein-tyrosine-kinase; primitive
KW hematopoietic cell; fetal liver kinase; diagnostic ligand
KW isolation; bone marrow disease therapy.
XX Homo sapiens.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide /note= "signal peptide"
FT Protein /note= "mature protein"
FT Domain /note= "extracellular receptor domain"
FT Domain /note= "transmembrane region"
FT Domain /note= "intracellular catalytic domain"
XX WO9500554-A.
XX 05-JAN-1995.
XX 17-JUN-1994; 94WO-US06944.

```
XX 18-JUN-1993; 93US-0080244.
PR 21-JUN-1993; 93US-0081508.
PR 23-NOV-1993; 93US-0157490.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Lemischka IR;
XX
DR WPI; 1995-052014/07.
DR N-PSDB; AAQ81013.
XX
PT Ligand for receptor protein tyrosine kinase - useful for the
PT stimulation of primitive haematopoietic stem cells causing
PT proliferation and/or differentiation
XX
PS Disclosure; Fig 1b; 131pp; English.
XX
CC The sequence corresponds to a human Flk2 (fetal liver kinase)
CC receptor protein-tyrosine-kinase, which is expressed in primitive
CC hematopoietic cells but not in mature hematopoietic cells. The
CC protein is useful in isolation of receptor ligands, which have
CC applications in diagnosis of bone marrow disorders and in
CC stimulating proliferation and/or differentiation of primitive
CC hematopoietic stem cells.
XX
SQ Sequence 993 AA;
Query Match 100.0%; Score 5274; DB 16; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVVSAMIFGTITNQDLPIKCVLINHKNNDSSVGKSSSPYMWSESP 60
DB 1 MPALARDAGTVPLLVVSAMIFGTITNQDLPIKCVLINHKNNDSSVGKSSSPYMWSESP 60
QY 61 EDLGCALRQSSGTVYAAAEVDVSASITLQVLVDPAGNISCLWFKHSSLCNQPHFDL 120
DB 61 EDLGCALRQSSGTVYAAAEVDVSASITLQVLVDPAGNISCLWFKHSSLCNQPHFDL 120
QY 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLTLRPPYRKMEQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLTLRPPYRKMEQD 180
QY 181 ALVCISESVPEIENVWLCDSQGESCKEESPAVVVKEEVLHGFQDTRCCARNELGRE 240
DB 181 ALVCISESVPEIENVWLCDSQGESCKEESPAVVVKEEVLHGFQDTRCCARNELGRE 240
QY 241 CTLFTIDLNOTPQTTLPQLFLKVGEPILTRCAVHVNHGFGLTWELENKALEEGNYFEM 300
DB 241 CTLFTIDLNOTPQTTLPQLFLKVGEPILTRCAVHVNHGFGLTWELENKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
QY 361 IDQYEEFCFSVREKAYPQIRCTWTFKSPCKEQGLDNGYSISKFCNHRKHQGEYIFHA 420
DB 361 IDQYEEFCFSVREKAYPQIRCTWTFKSPCKEQGLDNGYSISKFCNHRKHQGEYIFHA 420
QY 421 ENDDAOFKMTLNIRKPPQVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAOFKMTLNIRKPPQVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
QY 481 GVNRRKANRVFGQWSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GVNRRKANRVFGQWSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIWVLTLLCHKQKQRYESQLOMQVQSSDNEYFYVDPREYEYD 600
DB 541 NISFYATIGVCLLFIWVLTLLCHKQKQRYESQLOMQVQSSDNEYFYVDPREYEYD 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSREALMS 660
```

```
DB 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSREALMS 660
QY 661 ELKMMTQLGSHENIVNLGACTLSGPIYILFIFYCCYCDLLNLYLSRKREKPHRTWTEIFKE 720
DB 661 ELKMMTQLGSHENIVNLGACTLSGPIYILFIFYCCYCDLLNLYLSRKREKPHRTWTEIFKE 720
QY 721 HNFSEFPTQSHSPNSMPPGSRVQIHPDSQIISGLHGNPSHSEDEIEYENOKRLEEEDL 780
DB 721 HNFSEFPTQSHSPNSMPPGSRVQIHPDSQIISGLHGNPSHSEDEIEYENOKRLEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPPFYATEIYIIMOSCNWAFDSKRKPSFPNLTSLFGCQLADAEAMYNQVNDG 960
DB 901 LIQNGFKMDQPPFYATEIYIIMOSCNWAFDSKRKPSFPNLTSLFGCQLADAEAMYNQVNDG 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
RESULT 2
AAR67536
ID AAR67536 standard; Protein; 993 AA.
XX
AC AAR67536;
XX
DT 04-JUL-1995 (first entry)
XX
DE Human flk-2.
XX
KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Domain 28..543
FT Region 544..563
FT Domain 564..993
XX US5367057-A.
XX PN 22-NOV-1994.
XX
PF 02-APR-1991; 91US-0679666.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 19-NOV-1992; 92US-0977451.
XX 30-APR-1993; 93US-0055269.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1995-005894/01.
DR
```


DR N-PSDB: AAQ79069.
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
PT proliferation and/or stimulation of primitive mammalian
PT haematopoietic stem cells in vitro or in vivo.
XX
PS Disclosure; Fig. 2A-1F; 69pp; English.
XX
CC cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
CC respectively, and the deduced amino acid sequences in AAQ67533-37,
CC respectively.
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 16; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVVKVLINHKNDSSVKGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVVKVLINHKNDSSVKGKSSSYPMVSESP 60
QY 61 EDLGCALRQSSGTVYEAAREVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
Db 61 EDLGCALRQSSGTVYEAAREVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
QY 121 QNRGVSVWILKMTQTQAGYLLFIQSEATNTYTLFVSVTNTLLYTLRPPYFRKMENQD 180
Db 121 QNRGVSVWILKMTQTQAGYLLFIQSEATNTYTLFVSVTNTLLYTLRPPYFRKMENQD 180
QY 181 ALVCISESPEPTVEWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNEIGRE 240
Db 181 ALVCISESPEPTVEWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNEIGRE 240
QY 241 CTRLFIDLNQTPQTLPLQFLKVGEPILWIRKAVHVNHGFGFTWELENKALEBGNFYEM 300
Db 241 CTRLFIDLNQTPQTLPLQFLKVGEPILWIRKAVHVNHGFGFTWELENKALEBGNFYEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGTYTSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTGTYTSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYPOIRCTFTFSRKPCEQKGLDNGYSISKFCNHRKHQGEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTFTFSRKPCEQKGLDNGYSISKFCNHRKHQGEYIFHA 420
QY 421 ENDDAOFKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480
Db 421 ENDDAOFKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480
QY 481 GVNRRKANRVFGQWSSSTLNNSSEATKGLVKCCAYNSLGTSETILLNSPGPFPIQD 540
Db 481 GVNRRKANRVFGQWSSSTLNNSSEATKGLVKCCAYNSLGTSETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIIVTLILCHYKKQRYESQLOLVQVQVQSSDNEYFYDFREYED 600
Db 541 NISFYATIGVCLLFIIVTLILCHYKKQRYESQLOLVQVQVQSSDNEYFYDFREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGYTSIOAVKMLKEKADSSREALMS 660
Db 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGYTSIOAVKMLKEKADSSREALMS 660
QY 661 ELKMTQLGSHENIVNLLGACTLSGPTLYLIFEYCCYGDLLNLYLSKREKPHRTWTETPKE 720
Db 661 ELKMTQLGSHENIVNLLGACTLSGPTLYLIFEYCCYGDLLNLYLSKREKPHRTWTETPKE 720
QY 721 HNEFSFYPTFOSHNSPPGSRVQIHPDSQISGLHNSPHEDEIEYENQKRLREEDL 780
Db 721 HNEFSFYPTFOSHNSPPGSRVQIHPDSQISGLHNSPHEDEIEYENQKRLREEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIMSDS 840

Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWAPESLPEGYTTIKSDVMSYSGILLWEIFSLGVNYPGIPVDANEYK 900
Db 841 NYVVRGNARLPVKWAPESLPEGYTTIKSDVMSYSGILLWEIFSLGVNYPGIPVDANEYK 900
QY 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFLGCOLADAEAMYNQVNDG 960
Db 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFLGCOLADAEAMYNQVNDG 960
QY 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993
Db 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993

RESULT 3
AAR97419
ID AAR97419 standard; Protein; 993 AA.
XX
AC AAR97419;
XX
DT 11-DEC-1996 (first entry)
XX
DE Murine foetal liver kinase 2.
XX
DE Murine foetal liver kinase 2; flk-2; protein tyrosine kinase;
KW monoclonal; antibody; extracellular domain; receptor assay;
KW haematopoietic stem cell; ligand; stimulation; proliferation;
KW differentiation; treatment; anaemia; bone marrow damage;
KW cancer chemotherapy; radiation.
XX
QS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Peptide /label= sig_peptide
FT Peptide 28..993
FT Peptide /label= mat_peptide
FT Domain 28..544
FT Peptide /label= extracellular_domain
FT Domain 545..564
FT Domain /label= transmembrane_domain
FT Domain 565..993
FT Peptide /label= intracellular_domain
XX
PN US5548065-A.
XX
PD 20-AUG-1996.
XX
PF 02-APR-1991; 91US-0679666.
XX
PR 19-NOV-1992; 92US-0977451.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252517.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Lemischka IR;
XX
DR WPI; 1996-392678/39.
DR N-PSDB; AAT38734.
XX
PT Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
PT for isolating haematopoietic stem cells expressing receptor and for
PT obtaining ligands
XX
PS Claim 1; Columns 39-48; 50pp; English.

XX The present sequence is murine foetal liver kinase 2 (flk-2),
 CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-2 can be used to
 CC assay for flk receptors on the surface of primitive haematopoietic
 CC stem cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 XX
 XX Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 17; Length 993;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPPLLVFVSAMIFGTITNODLPVKVCLVNLHKNNDSSVGKSSSYPMVSESP 60
 DB 1 MPALARDAGTVPPLLVFVSAMIFGTITNODLPVKVCLVNLHKNNDSSVGKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTVEAAAVEVDVSITLQVLVDAPGNISCLWFKHSLNCPHFDL 120
 DB 61 EDLGCALRPOSSGTVEAAAVEVDVSITLQVLVDAPGNISCLWFKHSLNCPHFDL 120

QY 121 QNRGVVSMVTLKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPYFRKMNQD 180
 DB 121 QNRGVVSMVTLKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPYFRKMNQD 180

QY 181 ALVCISESVPETVWVLCDSQGESCKEESPAVVVKEEVLHFLGTDIRCCARNELGRE 240
 DB 181 ALVCISESVPETVWVLCDSQGESCKEESPAVVVKEEVLHFLGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNOTPOTLPQLFLKVGEPWIRKAVVHNGFGLTWELNKALEGNFEM 300
 DB 241 CTRLFTIDLNOTPOTLPQLFLKVGEPWIRKAVVHNGFGLTWELNKALEGNFEM 300

QY 301 STYSTNRTRIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIVGKGFNATNSSEDE 360
 DB 301 STYSTNRTRIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIVGKGFNATNSSEDE 360

QY 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKFCNKHQPGXEYIFHA 420
 DB 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKFCNKHQPGXEYIFHA 420

QY 421 ENDDAOTFKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTKKCDKSPNCTEITE 480
 DB 421 ENDDAOTFKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTKKCDKSPNCTEITE 480

QY 481 GVNRRKANRVFGOWVSSSTLNKSEAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
 DB 481 GVNRRKANRVFGOWVSSSTLNKSEAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIVVLTLLCHKQKQRYESQLOMVOVGTSSDNEYFYVDPREYEYD 600
 DB 541 NISFYATIGVCLLFIVVLTLLCHKQKQRYESQLOMVOVGTSSDNEYFYVDPREYEYD 600

QY 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLKEKADSSREALMS 660
 DB 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLKEKADSSREALMS 660

QY 661 ELKMTQLGSHENIVNLGACTLSGPIYLLFEYCCYGDLLNLYLSKREKPHRTWTEIFKE 720
 DB 661 ELKMTQLGSHENIVNLGACTLSGPIYLLFEYCCYGDLLNLYLSKREKPHRTWTEIFKE 720

QY 721 HNFSEFPTFOSHNSMPGSGREVQIHDPDSQISGLHGNHSHSEDEIEYENOKRLEEEEDL 780
 DB 721 HNFSEFPTFOSHNSMPGSGREVQIHDPDSQISGLHGNHSHSEDEIEYENOKRLEEEEDL 780

QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
 DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

QY 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 DB 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900

QY 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFGCOLADAEAMYNQYVDG 960
 DB 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFGCOLADAEAMYNQYVDG 960

QY 961 RVSECPHTYQNRPFESREMDLGLLSPOAQVEDS 993
 DB 961 RVSECPHTYQNRPFESREMDLGLLSPOAQVEDS 993

RESULT 4
 AAW19873
 ID AAW19873 standard; Protein; 993 AA.
 XX
 AC AAW19873;
 XX
 DT 19-AUG-1997 (first entry)
 XX
 DE Human flk-2 receptor.
 XX
 KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
 KW pTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT Protein /note= "Signal peptide"
 FT /note= "Mature flk-2"
 XX
 PN US5621090-A.
 XX
 PD 15-APR-1997.
 XX
 PF 02-APR-1991; 91US-0679666.
 XX
 PR 26-JUN-1992; 92US-0906397.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 WPI; 1997-235228/21.
 DR N-PSDB; AAT72117.
 XX
 PT Protein containing the extracellular domain of human flk-2 - used
 PT for identification of primitive haematopoietic cell proliferation
 PT and differentiation stimulatory ligands, e.g. for treating anaemia
 XX
 PS Claim 1; Fig 1B; 55pp; English.
 XX
 CC This sequence represents human fetal liver kinase 2 (flk2). flk-2 is
 CC a receptor protein tyrosine kinase (pTK) and is important in transducing
 CC putative self-renewal signals from the environment. flk-2 is expressed
 CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
 CC and it is thought that flk-2 is expressed in the entire primitive portion
 CC of the haematopoietic hierarchy. The invention concerns a recombinant
 CC nucleic acid, preferably mRNA, which encodes a protein containing only
 CC the extracellular domain of human flk-2 and lacking the flk-2 intra-
 CC cellular catalytic domain. The resultant protein represents a soluble
 CC form of flk-2 which is used to isolate specific ligands for flk-2. These
 CC ligands can be used to stimulate proliferation and/or differentiation of

CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused
CC by cancer treatment or radiation.
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 18; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARAGTVPLLVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSSPMWSESP 60
DB 1 MPALARAGTVPLLVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSSPMWSESP 60
QY 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
QY 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRPYFRKMNQD 180
DB 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRPYFRKMNQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESKESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESKESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTFLFTIDLNOTPTTLPLQLFLKVGEPWLTRCKAVHNVHGFGLTWELNKALEEGNFEM 300
DB 241 CTFLFTIDLNOTPTTLPLQLFLKVGEPWLTRCKAVHNVHGFGLTWELNKALEEGNFEM 300
QY 301 STYSTNRTRIRILFAVSSVARNDGTGYTCSKHPQSOSALVTIVGKGFNATNSSSEDE 360
DB 301 STYSTNRTRIRILFAVSSVARNDGTGYTCSKHPQSOSALVTIVGKGFNATNSSSEDE 360
QY 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFNHKKHQPGEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFNHKKHQPGEYIFHA 420
QY 421 ENDDAOFKMTINIRKPOVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAOFKMTINIRKPOVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
QY 481 GVMNRKANRKFQWVSSSTLNNSAIGFLVKCCAYNSLGTSETILLNSPGFPFIQD 540
DB 481 GVMNRKANRKFQWVSSSTLNNSAIGFLVKCCAYNSLGTSETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIIVVLLILCHYKQPRYESQLQMWQVGTGSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVVLLILCHYKQPRYESQLQMWQVGTGSSDNEYFYVDFREYED 600
QY 601 LWEPFRENLEFGKVLGSAFGKVMNATAGISKTGVSIOAVKMLKEKADSSEREAALS 660
DB 601 LWEPFRENLEFGKVLGSAFGKVMNATAGISKTGVSIOAVKMLKEKADSSEREAALS 660
QY 661 ELKMMTQOLGSHENIVNLGACTLGGPILLYFFCCYGDLLNLYRSKREKFRHTWTETFE 720
DB 661 ELKMMTQOLGSHENIVNLGACTLGGPILLYFFCCYGDLLNLYRSKREKFRHTWTETFE 720
QY 721 HNFSPYFTQSHPNMSPGSRVQIHPDSDQISGLHGNSFSEDEIEYENQKRLSEEDL 780
DB 721 HNFSPYFTQSHPNMSPGSRVQIHPDSDQISGLHGNSFSEDEIEYENQKRLSEEDL 780
QY 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHGKVKVTCDFGLARDINSDS 840
DB 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHGKVKVTCDFGLARDINSDS 840
QY 841 NVVVRNARLPVKWMAPESEIFGIYTIKSDVWSYGLLWEIFSLGVNPNYPGIPVDANFYK 900
DB 841 NVVVRNARLPVKWMAPESEIFGIYTIKSDVWSYGLLWEIFSLGVNPNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPPYATEEIIYIMOSWAFDSRKRPSFNLTSFLGCGCLADAEAMYQNVDG 960
DB 901 LIQNGFKMDQPPYATEEIIYIMOSWAFDSRKRPSFNLTSFLGCGCLADAEAMYQNVDG 960

QY 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993
DB 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993

RESULT 5

AAAY08617
ID AAY08617 standard; Protein; 993 AA.

XX AAY08617;
XX AC
XX XX
DT 05-AUG-1999 (first entry)
XX XX

DE Human flk-2 protein.

XX Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
KW monoclonal; polyclonal; antibody; tyrosine kinase.

XX Homo sapiens.

XX OS
XX PN US5912133-A.

XX PD 15-JUN-1999.

XX PF 10-FEB-1998; 98US-0021324.

XX PR 19-NOV-1992; 92US-0977451.

XX PR 02-APR-1991; 91US-0679666.

XX PR 28-JUN-1991; 91US-0728913.

XX PR 15-NOV-1991; 91US-0793065.

XX PR 24-DEC-1991; 91US-0813593.

XX PR 26-JUN-1992; 92US-0906397.

XX PR 12-NOV-1992; 92US-0975049.

XX PR 30-APR-1993; 93US-0055269.

XX PR 31-OCT-1994; 94US-0252498.

XX PR 15-FEB-1996; 96US-0601891.

XX (UYPR-) UNIV PRINCETON.

XX PI Lemischka IR;

XX WPI; 1999-357194/30.

XX N-PSDB; AAX77515.

XX Isolating hematopoietic cells expressing fetal liver kinase 1
PT receptors

XX PS Disclosure; Fig 1b; 59pp; English.

XX CC This invention describes a novel method of isolating cells expressing
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
CC binding the cells to a polyclonal or monoclonal antibody specific to
CC the flk-1 receptor and isolating the cells that have bound to the
CC antibody. The method can be used to isolate hematopoietic stem cells in
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
CC the invention belong to the receptor protein family. This sequence
CC represents the human flk-2 protein which is used in the method of the
CC invention.

XX SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 20; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARAGTVPLLVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSSPMWSESP 60

DB 1 MPALARAGTVPLLVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSSPMWSESP 60

QY 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

DB 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

```
QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQD 180
QY 181 ALVCISESPEPIVWLVCDQSGESCKEESPAVVVKEEKVLHFLGTDIRCCARNELGRE 240
DB 181 ALVCISESPEPIVWLVCDQSGESCKEESPAVVVKEEKVLHFLGTDIRCCARNELGRE 240
QY 241 CTRLFTIDLNTQPTTLPLQLFLKVGPELWTRCAVHVNHGFGTLWELENKALEEGNFEM 300
DB 241 CTRLFTIDLNTQPTTLPLQLFLKVGPELWTRCAVHVNHGFGTLWELENKALEEGNFEM 300
QY 301 STYSTNRWTRILFAFVSSVARNDGTGYTSSSKHPQSQSLVITVKGFTNATNSSDYE 360
DB 301 STYSTNRWTRILFAFVSSVARNDGTGYTSSSKHPQSQSLVITVKGFTNATNSSDYE 360
QY 361 IDQYEEFCFSVREKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGCEYIFHA 420
DB 361 IDQYEEFCFSVREKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGCEYIFHA 420
QY 421 ENDDAQTMTFTLNIRKPOVLAEASQASCFSDGYPLPSWTWKKCSKSPNCTEITE 480
DB 421 ENDDAQTMTFTLNIRKPOVLAEASQASCFSDGYPLPSWTWKKCSKSPNCTEITE 480
QY 481 GVNRRKANRVFCQWSSSTLNNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GVNRRKANRVFCQWSSSTLNNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLCHKHKQFRYESQLQMVQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIVVLTLLCHKHKQFRYESQLQMVQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOQAVKMLKEKADSSREALMS 660
DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOQAVKMLKEKADSSREALMS 660
QY 661 ELKMTQLGSHENIVNLGACTLSGPYILFEYCCYDGLNLYRSKREKPHRTWTETPKE 720
DB 661 ELKMTQLGSHENIVNLGACTLSGPYILFEYCCYDGLNLYRSKREKPHRTWTETPKE 720
QY 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780
DB 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780
QY 781 NVLTFFDLFCFAYQAVKGMFELEFKSCVHRDLAARNVLTHTGKVVKICDGLGARDIMSDS 840
DB 781 NVLTFFDLFCFAYQAVKGMFELEFKSCVHRDLAARNVLTHTGKVVKICDGLGARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWETFSLGVPNPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWETFSLGVPNPGIPVDANFYK 900
QY 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRKPSFPLNLSFGLCOLADAEAMYQNVDG 960
DB 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRKPSFPLNLSFGLCOLADAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPRFSEMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPRFSEMDLGLLSPQAQVEDS 993

RESULT 6
AAR37503
ID AAR37503 standard; Protein; 993 AA.
XX
AC AAR37503;
XX
DE 19-OCT-1993 (first entry)
XX
DE Human flk-2.
XX
KW Murine; receptor; protein; tyrosine kinase; ptk; flk-2; primitive;
```

```
KW hematopoietic cell; mature; family; conserved; region;
KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;
KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;
KW hierarchy; transduction; T-lymphoid; lineage.
XX
OS Homo sapiens.
XX
FH Key
XX Location/Qualifiers
XX 1..27
XX /note= "Hydrophobic leader sequence"
FT Peptide
XX /note= "Mature murine flk-2"
FT Protein
XX 28..993
FT Domain
XX 28..543
FT /note= "Extracellular receptor domain"
FT Region
XX 544..563
FT /note= "Transmembrane region"
FT Domain
XX 564..983
FT /note= "Intracellular catalytic domain"
XX
XX WO9310136-A.
XX
XX 27-MAY-1993.
XX
XX 16-NOV-1992; 92WO-US09893.
XX
XX 15-NOV-1991; 91US-0793065.
XX
XX (UYP- ) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1993-182479/22.
XX
XX N-PSDB; AAQ40915.
XX
XX Totipotent haematopoietic stem cell receptors, their ligands and
XX DNA sequences - for treating anaemia(s) and bone marrow damage
XX due to e.g. cancer chemotherapy or radiotherapy
XX
XX Claim 39; Fig 1b; 127pp; English.
XX
XX This sequence represents the human receptor protein tyrosine kinase
XX (PTK), flk-2. The nucleic acid encoding this receptor is expressed
XX in primitive hematopoietic cells and not in mature hematopoietic
XX cells. Members of this family of ptk's can be recognised by the
XX conserved amino acid regions in the catalytic domain. This family
XX of ptk's also contains c-kit. These new receptors are termed fetal
XX liver kinases (flk's) after the tissue in which they were discovered.
XX flk-2 is also expressed in fetal spleen, fetal thymus, adult brain
XX and adult bone marrow. flk-2 is expressed in individual multipotential
XX CFU-Blast colonies capable of generating numerous multilineage colonies
XX upon replating. It is likely therefore, that flk-2 is expressed in
XX the entire primitive portion of the hematopoietic hierarchy. This is
XX consistent with flk-2 being important in transducing putative self-
XX renewal signals from the environment. flk-2 is the first receptor
XX ptk known to be expressed in the T-lymphoid lineage.
XX
XX Sequence 993 AA;
```

```
Query Match 99.9%; Score 5271; DB 14; Length 993;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 992; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVIRKVLINHKNNDSVKGSSYPMVSESP 60
DB 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVIRKVLINHKNNDSVKGSSYPMVSESP 60
QY 61 EDLGCALRPQSSGTVYEAADVDSASITLQVLVDAPGNTSCLWFKHSSLNCOPHFDL 120
DB 61 EDLGCALRPQSSGTVYEAADVDSASITLQVLVDAPGNTSCLWFKHSSLNCOPHFDL 120
QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQD 180
```

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHLEFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHLEFGTDIRCCARNELGRE 240
QY 241 CTRLTIDLNQTPQTTLPQLFLKVGEPILWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
DB 241 CTRLTIDLNQTPQTTLPQLFLKVGEPILWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
QY 301 STYSTNRMTIRILFAFVSSVARNDTGYTSCSSKHPQSQSLVITVIGKGFNATNSSDEYE 360
DB 301 STYSTNRMTIRILFAFVSSVARNDTGYTSCSSKHPQSQSLVITVIGKGFNATNSSDEYE 360
QY 361 IDOYEFCFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHRKHOPGEYIFHA 420
DB 361 IDOYEFCFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHRKHOPGEYIFHA 420
QY 421 ENDDAQFTKMTILNIRKPOVLAEASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAQFTKMTILNIRKPOVLAEASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
QY 481 GWNRRKANRVFQWVSSSTLNNSAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GWNRRKANRVFQWVSSSTLNNSAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIYVLTLLCHKKYKQRYESOLQMVQVGTSSDNEYFYVDREYEYD 600
DB 541 NISFYATIGVCLLFIYVLTLLCHKKYKQRYESOLQMVQVGTSSDNEYFYVDREYEYD 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATYIGISKTGVSIOQVAVKMLKEKADSSERIALMS 660
DB 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATYIGISKTGVSIOQVAVKMLKEKADSSERIALMS 660
QY 661 ELKMTQLGSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYRSRKEKPHRTWTETIFE 720
DB 661 ELKMTQLGSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYRSRKEKPHRTWTETIFE 720
QY 721 HNFSEFPTQSHNSMPSGREGVQIHPDSQIISGLHGNSFHSDEIEYENOKRLEEEEDL 780
DB 721 HNFSEFPTQSHNSMPSGREGVQIHPDSQIISGLHGNSFHSDEIEYENOKRLEEEEDL 780
QY 781 NVLTFFDLCLCFAYQAKGMEFFLFCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840
DB 781 NVLTFFDLCLCFAYQAKGMEFFLFCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840
QY 841 NYVVRNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPYGPVPDANFYK 900
DB 841 NYVVRNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPYGPVPDANFYK 900
QY 901 LIONGFKMDOPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFGCQLADAEAMYQNVDG 960
DB 901 LIONGFKMDOPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFGCQLADAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993
RESULT 7
AAR44995
ID AAR44995 standard; Protein; 993 AA.
XX
AC AAR44995;
XX
DT 27-JUN-1994 (first entry)
XX
DE Human flk-2 receptor protein tyrosine kinase.
XX
KW Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
KW hflx; primitive; totipotent; haematopoietic cell; stem cell;
KW proliferation; stromal cell.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..27 /label= signal_sequence
FT /note= "hydrophobic leader"
FT 28..993 /label= flk-2
FT /note= "mature protein"
FT 28..544 /label= extracellular_domain
FT 545..563 /label= transmembrane_region
FT 564..993 /label= intracellular_catalytic_domain
XX
PN US5270458-A.
XX
PD 14-DEC-1993.
XX
PF 02-APR-1991; 91US-0679666.
XX
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 19-NOV-1992; 92US-0977451.
XX
(UUPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
DR WP1; 1993-405021/50.
DR N-PSDB; AAQ53503.
XX
PT Isolated nucleic acid molecules of hematopoietic stem cell
PT receptor flk-2 - encoding mammalian receptor protein tyrosine
PT kinases expressed in primitive haematopoietic cells
XX
PS Disclosure; Fig 1b; 60pp; English.
XX
CC Nucleic acid sequences coding for murine flk-2 and specified
CC subfragments of it are claimed. The human flk-2 coding sequence
CC (i.e. AAQ53503) is also disclosed. The murine and human flk-2
CC polypeptides represent a new class of receptor protein tyrosine
CC kinases which are expressed only in primitive haematopoietic cells.
XX
SQ Sequence 993 AA;
Query Match 99.8%; Score 5266; DB 14; Length 993;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVFSAIFGTITNODLPVIRKVLINHKNDSSVSKSSYPMVSESP 60
DB 1 MPALARDAGTVPLLVFSAIFGTITNODLPVIRKVLINHKNDSSVSKSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVEEAAAEVDVDSASITLQVLVDAPGNISCLWFKHSSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVEEAAAEVDVDSASITLQVLVDAPGNISCLWFKHSSLNCQPHFDL 120
QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHLEFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHLEFGTDIRCCARNELGRE 240
QY 241 CTRLTIDLNQTPQTTLPQLFLKVGEPILWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
DB 241 CTRLTIDLNQTPQTTLPQLFLKVGEPILWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300

Qy	301	STYSTNRTMIRILFAPVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDYE	360
Db	301	STYSTNRTMIRILFAPVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDYE	360
Qy	361	IDQEEFCFSVRKAYPQIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHQPGGEYIFHA	420
Db	361	IDQEEFCFSVRKAYPQIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHQPGGEYIFHA	420
Qy	421	ENDDAQTKMFTLNIRRKPOVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTBEITE	480
Db	421	ENDDAQTKMFTLNIRRKPOVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTBEITE	480
Qy	481	GVNRRKANRKYGVQWVSSSTLNNSEAIKGLYKCCAYNSLGTSCETILLNSPGPEFIQD	540
Db	481	GVNRRKANRKYGVQWVSSSTLNNSEAIKGLYKCCAYNSLGTSCETILLNSAGFFFIQD	540
Qy	541	NTSFYATIGVCLLFIIVVLTLLICHKKYKFRYESQIQMVQVQVGTSSDNEYFYVDFREYED	600
Db	541	NTSFYATIGVCLLFIIVVLTLLICHKKYKFRYESQIQMVQVQVGTSSDNEYFYVDFREYED	600
Qy	601	LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREA	660
Db	601	LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREA	660
Qy	661	ELKMMTQLGSHENIVNLLGACTLSGPIYLIFBYCCYGDLLNLYLSRREKPHRTWTET	720
Db	661	ELKMMTQLGSHENIVNLLGACTLSGPIYLIFBYCCYGDLLNLYLSRREKPHRTWTET	720
Qy	721	HNFSFYPTQSHNSMPGSRVQIHPDSDQISGLHGNSEFHSDETEYENOKRLEEBE	780
Db	721	HNFSFYPTQSHNSMPGSRVQIHPDSDQISGLHGNSEFHSDETEYENOKRLEEBE	780
Qy	781	NVLTFEDLLCFAYQVAKGMEFLEKFCVHRDLAARNVLTTHGKVKIKCDFGLARDIMS	840
Db	781	NVLTFEDLLCFAYQVAKGMEFLEKFCVHRDLAARNVLTTHGKVKIKCDFGLARDIMS	840
Qy	841	NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYGILLWEIFSLGVNPNYPGIPVD	900
Db	841	NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYGILLWEIFSLGVNPNYPGIPVD	900
Qy	901	LIONGFKMDOPFYATEIYIIMOSWAFDSRKRPSFNLTSFLGCOLADAEAEAMYQNV	960
Db	901	LIONGFKMDOPFYATEIYIIMOSWAFDSRKRPSFNLTSFLGCOLADAEAEAMYQNV	960
Qy	961	RVSECPHTYQNRPRPFGRMDLGLLSPOAQVEDS	993
Db	961	RVSECPHTYQNRPRPFGRMDLGLLSPOAQVEDS	993
RESULT 8			
AAR75961			
ID	AAR75961 standard; Protein; 993 AA.		
XX			
AC	AAR75961;		
XX			
XX	29-DEC-1995 (first entry)		
DT	Human STK-1.		
XX			
XX	STK-1; receptor PTK; protein tyrosine kinase.		
XX	Homo sapiens.		
OS			
FX	Key	Location/Qualifiers	
FX	Peptide	1..23	
FT	/label= signal		
FT	/note= "unique PTK receptor sequence"		
FT	Cleavage-site	25..26	
FT	Domain	35..524	
FT	/note= "Part of extracellular domain contg. 22 Cys residues marked on Fig. 1"		
FT	Modified-site	43..45	

FT	Modified-site	/label= potential N-linked glycosylation site
FT	Modified-site	100..102
FT	Modified-site	/label= see above
FT	Modified-site	151..153
FT	Modified-site	/label= see above
FT	Modified-site	250..252
FT	Modified-site	/label= see above
FT	Modified-site	306..308
FT	Modified-site	/label= see above
FT	Modified-site	323..325
FT	Modified-site	/label= see above
FT	Modified-site	351..353
FT	Modified-site	/label= see above
FT	Modified-site	354..356
FT	Modified-site	/label= see above
FT	Modified-site	473..475
FT	Modified-site	/label= see above
FT	Modified-site	502..504
FT	Modified-site	/label= see above
FT	Region	542..562
FT	Region	/label= transmembrane spanning region
FT	Region	708..782
FT	Domain	/label= kinase insert region
FT	Domain	617..622
FT	Domain	/label= ATP binding domain
FT	Domain	835..840
FT	Domain	/label= WMAPES motifs
FT	Peptide	/note= "cytoplasmic domain"
FT	Peptide	808..813
FT	Peptide	/note= "used to design PCR oligos"
FT	Peptide	870..875
FT	Peptide	/note= "used to design PCR oligos"
XX		
PN		W09519175-A.
XX		
XX		20-JUL-1995.
PD		
XX		
XX		06-JAN-1995; 95WO-US00176.
XX		
XX		14-JAN-1994; 94US-0183211.
XX		
PR		(UYJO) UNIV JOHNS HOPKINS.
PA		(UYPE-) UNIV PENNSYLVANIA.
XX		
PI		Civin CI, Gewirtz AM, Small D;
XX		
DR		WPI; 1995-263709/34.
DR		N-PSDB; AAQ91356.
XX		
XX		Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide
PT		used to treat neoplastic diseases and as bone marrow purging
PT		agents for treating leukaemia and neoplasma
PT		
XX		
PS		Disclosure; Fig 1; 66pp; English.
XX		
CC		The STK-1 gene encodes a receptor PTK which is expressed in
CC		proliferating hematopoietic stem cells but not in quiescent stem
CC		cells. The STK-1 gene is also expressed in certain malignant
CC		cells of non-hematopoietic origin. An antisense oligo specific for
CC		STK-1 is an oligo having a sequence (i) capable of forming a stable
CC		triplex with a portion of the STK-1 gene, or (ii) capable of forming
CC		a stable duplex with a portion of an mRNA transcript of the STK-1
CC		gene. Antisense oligos capable of forming a stable duplex with a
CC		portion of a STK-1 mRNA transcript are given in AAQ91536 FT and in
CC		AAQ91537 and AAQ91538. The antisense oligos of the invention are useful
CC		in the treatment of hematologic malignancies characterised by
CC		STK-1 expression. Several of the conserved domains of PTKs
CC		including the ATP binding domain and the WMAPES motifs are
CC		found in the STK-1 protein (see AAR75961 FT).
XX		
SQ	Sequence	993 AA;
Query Match		99.8%; Score 5266; DB 16; Length 993;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPDLLVFSAMIFGTITNODLPVVKVLIHNKNDSSVYKSSSYPMVSESP 60
DB 1 MPALARDAGTVPDLLVFSAMIFGTITNODLPVVKVLIHNKNDSSVYKSSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWFKHSLNCPHF 120
DB 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWFKHSLNCPHF 120
QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNQTTQTLPLQFLKVGEPDLWIRKAVHVNHGFGLTWELNKALEBGNFEM 300
DB 241 CTRFLTIDLNQTTQTLPLQFLKVGEPDLWIRKAVHVNHGFGLTWELNKALEBGNFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATNSSEDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATNSSEDEYE 360
QY 361 IDQYEFECFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
DB 361 IDQYEFECFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
QY 421 ENDDAOFKMTINIRKPOVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAOFKMTINIRKPOVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
QY 481 GVMNRKANRVQWYSSSTLNNSAIEKGLVKCCAYNSLGTSCETILLNSPFPPIQD 540
DB 481 GVMNRKANRVQWYSSSTLNNSAIEKGLVKCCAYNSLGTSCETILLNSPFPPIQD 540
QY 541 NTSFYATIGVCLLFIVVLLILCHYKQPRYESQLOMVQVTSQSSDNEYFYVDFREYEYD 600
DB 541 NTSFYATIGVCLLFIVVLLILCHYKQPRYESQLOMVQVTSQSSDNEYFYVDFREYEYD 600
QY 601 LKWEFPRENLEFGKVLGSAFGVMKATAYGISKTVGSIOVAVKMLKEKADSSEREA 660
DB 601 LKWEFPRENLEFGKVLGSAFGVMKATAYGISKTVGSIOVAVKMLKEKADSSEREA 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPTIYLFYCCYGDLLNLYRSKREKPHRTWTETFE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPTIYLFYCCYGDLLNLYRSKREKPHRTWTETFE 720
QY 721 HNFSPYPTQSHPNSSMPGSRVQIHPDSDQISGLHGNPSHSEDEIYENOKRLEEBEDL 780
DB 721 HNFSPYPTQSHPNSSMPGSRVQIHPDSDQISGLHGNPSHSEDEIYENOKRLEEBEDL 780
QY 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHKVVATCDPGLARDINSDS 840
DB 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHKVVATCDPGLARDINSDS 840
QY 841 NVVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLMEIFSLGVNYPYGPVPDANFYK 900
DB 841 NVVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLMEIFSLGVNYPYGPVPDANFYK 900
QY 901 LQNGFKMOPPYATETIYIIQSWAFPSRKRPSNLTSLFGCOLADEEAMYNQVNDG 960
DB 901 LQNGFKMOPPYATETIYIIQSWAFPSRKRPSNLTSLFGCOLADEEAMYNQVNDG 960
QY 961 RVSECPHTTQNRPPFSREMDLGLLSPOAQVEDS 993
DB 961 RVSECPHTTQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 9

AAR81869
ID AAR81869 standard; Protein; 993 AA.

XX
AC AAR81869;

XX
DT 19-MAR-1996 (first entry)

XX
DE Human Flk2/flt3 tyrosine kinase receptor.

XX
KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;
KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell.
XX
OS Homo sapiens.

XX
PN W09527062-A1.

XX
PD 12-OCT-1995.

XX
PF 23-MAR-1995; 95WO-US03718.

XX
PR 04-APR-1994; 94US-0222299.

XX
PA (GETH) GENENTECH INC.

XX
PI Bennett BD, Broz SD, Matthews W, Zeigler FC;

XX
DR WPI; 1995-358636/46.

XX
DR N-PSDB; AAT00802.

XX
PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
PT - enhances proliferation of haematopoietic stem cells, in the
PT treatment of hypoplasia, anaemia, etc.

XX
PS Disclosure; Page 44-47; 59pp; English.

XX
CC DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2
CC (also called flt3) (AAR81868) was obtd. by RT-PCR amplification of RNA
CC isolated from mid-gestation mouse foetal livers using primers based
CC on the murine flt3 sequence, and subcloning of the product into pRK5.1.
CC An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was
CC constructed and used to raise agonist antibodies able to bind to, and
CC activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and
CC amino acid sequence (AAR81869) are also given.

XX
SQ Sequence 993 AA;

Query Match 99.8%; Score 5265; DB 16; Length 993;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPDLLVFSAMIFGTITNODLPVVKVLIHNKNDSSVYKSSSYPMVSESP 60

DB 1 MPALARDAGTVPDLLVFSAMIFGTITNODLPVVKVLIHNKNDSSVYKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWFKHSLNCPHF 120

DB 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWFKHSLNCPHF 120

QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180

DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180

QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240

DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240

QY 241 CTRFLTIDLNQTTQTLPLQFLKVGEPDLWIRKAVHVNHGFGLTWELNKALEBGNFEM 300

DB 241 CTRFLTIDLNQTTQTLPLQFLKVGEPDLWIRKAVHVNHGFGLTWELNKALEBGNFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATNSSEDEYE 360

DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATNSSEDEYE 360


```
QY 361 IDQYEFCEFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKGQGEYIFHA 420
DB 361 IDQYEFCEFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKGQGEYIFHA 420
QY 421 ENDDAQFTKMFNLNRRKPKVLAESAASQASCFSDGYPPLSPWTWKCKSDKSPNCTEEITE 480
DB 421 ENDDAQFTKMFNLNRRKPKVLAESAASQASCFSDGYPPLSPWTWKCKSDKSPNCTEEITE 480
QY 481 GWNKARKKVFQGVSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GWNKARKKVFQGVSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NLSFYATIGVCLLFIWVLLCHKVKKFRYESQLQMVQVGTSSDNEYFYVDFREVEYD 600
DB 541 NLSFYATIGVCLLFIWVLLCHKVKKFRYESQLQMVQVGTSSDNEYFYVDFREVEYD 600
QY 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGQAVKMLKEKADSSERREALMS 660
DB 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGQAVKMLKEKADSSERREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFECYCGDLLNLYLSRKREKPHRTWTTEIFKE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFECYCGDLLNLYLSRKREKPHRTWTTEIFKE 720
QY 721 HNFSEYPTTQSHPNSSMPGSRVQIHPDSDQISLGHGNSFHSDELEYENQKRLKEEEDL 780
DB 721 HNFSEYPTTQSHPNSSMPGSRVQIHPDSDQISLGHGNSFHSDELEYENQKRLKEEEDL 780
QY 781 NVLTFFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIMSDS 840
DB 781 NVLTFFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPYGPVDFANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPYGPVDFANFYK 900
QY 901 LQNGFKMQDPFVATIEIYIMOSWAFDSKRPPNLTSLFGLCOLADAEAMYQNVDG 960
DB 901 LQNGFKMQDPFVATIEIYIMOSWAFDSKRPPNLTSLFGLCOLADAEAMYQNVDG 960
QY 961 RVSECPHTYONRRPFSREMDLGLLSPOAIVEDS 993
DB 961 RVSECPHTYONRRPFSREMDLGLLSPOAIVEDS 993

RESULT 10
AAR31376
ID AAR31376 standard; Protein; 1167 AA.
XX AC AAR31376;
XX DT 25-JUN-1993 (first entry)
XX DE Human flk-2.
XX KW Human; receptor; protein; tyrosine kinase; ptk; primitive; mammalian;
KW hematopoietic cell; ptk; mature; mhc; fetal; liver kinase 2; flk-2;
KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;
KW multipotent; T-lymphoid; lineage.
XX OS Homo sapiens.
XX FH Key
XX FT Misc-difference 1013
FT FT Location/Qualifiers
FT FT /note= "Nonsense mutation"
FT FT Misc-difference 1038
FT FT /note= "Nonsense mutation"
FT FT Misc-difference 1042
FT FT /note= "Nonsense mutation"
FT FT Misc-difference 1057
FT FT /note= "Nonsense mutation"
FT FT Misc-difference 1102
```

```
FT Misc-difference 1117 /note= "Nonsense mutation"
FT /note= "Nonsense mutation"
FT Misc-difference 1127 /note= "Nonsense mutation"
XX WO9300349-A.
XX PN
XX PD 07-JAN-1993.
XX 26-JUN-1992; 92WO-US05401.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 02-APR-1992; 92WO-US02750.
XX (UYPR-) UNIV PRINCETON.
XX PI Lemischka IR;
XX WPI: 1993-036323/04.
XX N-PSDB; AAQ35250.
XX Nucleic acid encoding receptor protein tyrosine kinase - allows
XX development of ligands to stimulate proliferation and/or
XX differentiation of mammalian haematopoietic stem cells
XX Claim 9; Fig 1b; 78pp; English.
XX This sequence represents a human receptor protein tyrosine kinase
XX which belongs to a new functional class of protein tyrosine kinases
XX (ptks). ptk's in this class are expressed in primitive mammalian
XX hematopoietic (phc) cells but not in mature hematopoietic cells (mhc).
XX This protein is an example of a receptor ptk and is called fetal liver
XX kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and thymus,
XX and adult brain and marrow. Expression of flk-2 mRNA occurs in the
XX most primitive thymocyte subset, which is believed to be uncommitted.
XX Therefore, thymocytes expressing flk-2 may be multipotential. flk-2
XX is the first receptor tyrosine kinase known to be expressed in the
XX T-lymphoid lineage.
XX SQ Sequence 1167 AA;
XX Query Match 99.8%; Score 5262; DB 14; Length 1167;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVFSAMIFGTITNQDLPVICKVLINHNKNDSSVGKSSYPMVSESP 60
DB 20 MPALARDAGTVPLLVFSAMIFGTITNQDLPVICKVLINHNKNDSSVGKSSYPMVSESP 79
QY 61 EDLGCALRPQSSGTVYERAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120
DB 80 EDLGCALRPQSSGTVYERAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 139
QY 121 QNRGVSVNVLKMTETQAGEYLLFTQSEATNYTILFTVSIINTLLYTLRRPYFRKMNQD 180
DB 140 QNRGVSVNVLKMTETQAGEYLLFTQSEATNYTILFTVSIINTLLYTLRRPYFRKMNQD 199
QY 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCARNELGRE 240
DB 200 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCARNELGRE 259
QY 241 CTRLFTIDLNOTPOTTLPOLEFLKVGGEPLWIRCKAVVNHGFGLTWELENKALEEGNYFEM 300
DB 260 CTRLFTIDLNOTPOTTLPOLEFLKVGGEPLWIRCKAVVNHGFGLTWELENKALEEGNYFEM 319
QY 301 STYSTNRTMIRILFAFYSSVARNDTGYTCSKSHPSOSALVTIVGKGFINATNSEDYE 360
DB 320 STYSTNRTMIRILFAFYSSVARNDTGYTCSKSHPSOSALVTIVGKGFINATNSEDYE 379
QY 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKGQGEYIFHA 420
```


QY 651 DSSREALMSELKMMTQGLSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKREKF 710
DB 661 DSSREALMSELKMMTQGLSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKREKF 720
QY 711 HRTWTEIFKEHNFSPYTFQSHPNSSMPGSRREVQIHPDSQDQISGLHGSFHSDEDEIYEN 770
DB 721 HRTWTEIFKEHNFSPYTFQSHPNSSMPGSRREVQIHPDSQDQISGLHGSFHSDEDEIYEN 780
QY 771 QKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDF 830
DB 781 QKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDF 840
QY 831 GLARDIMSDSNVVRGNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGWNYP 890
DB 841 GLARDIMSDSNVVRGNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGWNYP 900
QY 891 GIPVDANFYKLIQNGFKMDQFPYATEEYIIMQSCWAFDSRRPFPNLTSLGCOLADA 950
DB 901 GIPVDANFYKLIQNGFKMDQFPYATEEYIIMQSCWAFDSRRPFPNLTSLGCOLADA 960
QY 951 EAMQYQNVGRVSECPHTYQNR 973
DB 961 EAMQYQNVGRVSECPHTYQNR 983

RESULT 12

AAW63587
ID AAW63587 standard; Protein; 986 AA.
AC AAW63587;
XX
XX
DT 12-OCT-1998 (first entry)
XX
XX
DE Human receptor type protein kinase FLT3 protein SEQ ID NO:18.
KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.
XX
XX Homo sapiens.
PN WO9817808-A1.
XX
XX 30-APR-1998.
XX
XX 13-OCT-1997; 97WO-JP03667.
XX
XX 18-OCT-1996; 96JP-0297329.
XX
XX (TAKI) TAKARA SHUZO CO LTD.
XX
XX Yokota S;
PI
XX WPI; 1998-362333/31.
XX N-PSDB; AAV39040.
XX
PT Nucleic acid sequences encoding receptor type protein kinase -
PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
M2
XX
XX Claim 6; Page 40-45; 80pp; Japanese.

CC New nucleic acid sequences have been isolated which encode receptor type
CC protein kinases (especially a tyrosine kinases) having tandem repeats in
CC the juxtamembrane region. Also described in the present invention are:
CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589),
CC representing FLT3 juxtamembrane receptor type protein kinases found in
CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences
CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
CC recognising the kinases or their portions including the tandem repeat
CC region; (3) nucleic acid sequences hybridising with the nucleic acid;
CC (4) a method for detecting the nucleic acid sequences in human tissue
CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
CC gene, and (iii) determining the size of the gene for comparing the size

CC of the normal gene not containing tandem repeats, and (5) kits for
CC carrying out the detection. The products and methods may be used for
CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.

XX Sequence 986 AA;

QY Query Match 97.8%; Score 5157.5; DB 19; Length 986;
DB Best Local Similarity 98.7%; Pred. No. 0;
Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPALARDAGTVPVLLVYFVSAMIFGTITNQDLPVTKVLINHNKNDSSVGKSSSYPMVSESP 60
DB 1 MPALARDAGTVPVLLVYFVSAMIFGTITNQDLPVTKVLINHNKNDSSVGKSSSYPMVSESP 60

QY 61 EDLGCALRQSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLCNCQPHFDL 120
DB 61 EDLGCALRQSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLCNCQPHFDL 120

QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTILFTVTSIRNTLLYTLRRPYFRKMNQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTILFTVTSIRNTLLYTLRRPYFRKMNQD 180

QY 181 ALVCISESVPEIPEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEIPEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

QY 241 CTRFLTIDLNOTPQTTLPOLFLKVGPELWIRKAVHVNHGFLTWELNKALEGNFEM 300
DB 241 CTRFLTIDLNOTPQTTLPOLFLKVGPELWIRKAVHVNHGFLTWELNKALEGNFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGYTCTSSSKHPSQSALVTIIVGKGFNATNSSEDE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGYTCTSSSKHPSQSALVTIIVGKGFNATNSSEDE 360

QY 361 IDOYEFCFSVRKAYPQIRCTWTFSRKSPPCEQKGLDNGYSISKFNHKKHQGEYIFHA 420
DB 361 IDOYEFCFSVRKAYPQIRCTWTFSRKSPPCEQKGLDNGYSISKFNHKKHQGEYIFHA 420

QY 421 ENDDAOTFKMTLINIRKPOVLAEASASQSCFSDGYPLPSWTWKKCSDKSPNCTEITE 480
DB 421 ENDDAOTFKMTLINIRKPOVLAEASASQSCFSDGYPLPSWTWKKCSDKSPNCTEITE 480

QY 481 GVNRRKANRVFGQWVSSSTLNMSEAIKGLVKCCAYNSLTSCETILLNSPGPFPIQD 540
DB 481 GVNRRKANRVFGQWVSSSTLNMSEAIKGLVKCCAYNSLTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIVVLTLLICHKKYKQFYESQLOMVQVGTGSSDNEYFYV----- 592
DB 541 NISFYATIGVCLLFIVVLTLLICHKKYKQFYESQLOMVQVGTGSSDNEYFYVDFREYED 600

QY 593 -----DFREYEDLKWEFFPRENLEFGKVGSGAFKGVNNATAYGISTGVSIOVAVKMLK 647
DB 601 LKWEFFPREYEDLKWEFFPRENLEFGKVGSGAFKGVNNATAYGISTGVSIOVAVKMLK 660

QY 648 EKADSSREALMSELKMMTQGLSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKR 707
DB 661 EKADSSREALMSELKMMTQGLSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKR 720

QY 708 EKPHRTWTEIFKEHNFSPYTFQSHPNSSMPGSRREVQIHPDSQDQISGLHGSFHSDEIE 767
DB 721 EKPHRTWTEIFKEHNFSPYTFQSHPNSSMPGSRREVQIHPDSQDQISGLHGSFHSDEIE 780

QY 768 YENQKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKI 827
DB 781 YENQKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKI 840

QY 828 CDEGLARDIMSDSNVVRGNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVN 887
DB 841 CDEGLARDIMSDSNVVRGNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVN 900

QY 888 PYPGIPVDANFYKLIQNGFKMDQFPYATEEYIIMQSCWAFDSRRPFPNLTSLGCOL 947
DB 901 PYPGIPVDANFYKLIQNGFKMDQFPYATEEYIIMQSCWAFDSRRPFPNLTSLGCOL 960

Qy 948 ADAEAMYNQVNDGRVSECPHTYQNR 973
 Db 961 ADAEAMYNQVNDGRVSECPHTYQNR 986

RESULT 13

AAW63589
 ID AAW63589 standard; Protein; 986 AA.

XX AC AAW63589;

XX DT 12-OCT-1998 (first entry)

XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:20.

XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
 XX KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.

XX OS Homo sapiens.

XX PN WO9817808-A1.

XX PD 30-APR-1998.

XX PF 13-OCT-1997; 97WO-JP03667.

XX PR 18-OCT-1996; 96JP-0297329.

XX PA (TAKI) TAKARA SHUZO CO LTD.

XX PI Yokota S;

XX DR WPI; 1998-362333/31.

XX DR N-PSDB; AAV39042.

XX PT Nucleic acid sequences encoding receptor type protein kinase -
 PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
 PT M2

XX PS Claim 6; Page 50-55; 80pp; Japanese.

XX CC New nucleic acid sequences have been isolated which encode receptor type
 CC protein kinases (especially a tyrosine kinases) having tandem repeats in
 CC the juxtamembrane region. Also described in the present invention are:
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)
 CC representing FLT3 juxtamembrane receptor type protein kinases found in
 CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences
 CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
 CC recognising the kinases or their portions including the tandem repeat
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;
 CC (4) a method for detecting the nucleic acid sequences in human tissue
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
 CC gene, and (iii) determining the size of the gene for comparing the size
 CC of the normal gene not containing tandem repeats; and (5) kits for
 CC carrying out the detection. The products and methods may be used for
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.

XX Sequence 986 AA;

Query Match 97.8%; Score 5157.5; DB 19; Length 986;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

Qy 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60

Db 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTVYEAARAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120

Db 61 EDLGCALRPOSSGTVYEAARAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120

Qy 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180

Db 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180
 Qy 181 ALVCISESVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHFLFTDTRCCARNELGRE 240
 Db 181 ALVCISESVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHFLFTDTRCCARNELGRE 240
 Qy 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300
 Db 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300
 Qy 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
 Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
 Qy 361 IDQYEEFCFSVRFKAYPQIRCTWTSRKSFPCEQKGLDNQYSISKFCNKHQHPGEYIFHA 420
 Db 361 IDQYEEFCFSVRFKAYPQIRCTWTSRKSFPCEQKGLDNQYSISKFCNKHQHPGEYIFHA 420
 Qy 421 ENDDAQFTKMTFLNIRRRKQVLAELASASQSCFSDGYPLPSWTWKKCSDKSPNCTEITE 480
 Db 421 ENDDAQFTKMTFLNIRRRKQVLAELASASQSCFSDGYPLPSWTWKKCSDKSPNCTEITE 480
 Qy 481 GYWNRKANRKVFQGVQVSSSTLNKSEAIKFLVKCCAYNSLGTSCETILLNSPGPPFIOD 540
 Db 481 GYWNRKANRKVFQGVQVSSSTLNKSEAIKFLVKCCAYNSLGTSCETILLNSPGPPFIOD 540
 Qy 541 NISFYATIGVCLLFIVVLTLLICHKYKKQFYESQLQMVQVT-----GSSDN 587
 Db 541 NISFYATIGVCLLFIVVLTLLICHKYKKQFYESQLQMVQVT-----GSSDN 587
 Qy 588 EYFYVDFREYEDLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIQAVKMLK 647
 Db 601 EYFYVDFREYEDLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIQAVKMLK 660
 Qy 648 EKADSSEREALMSLKMMTQLGSHENIVNLGACTLSGPIYLIFECYCGDLLNLYRSKR 707
 Db 661 EKADSSEREALMSLKMMTQLGSHENIVNLGACTLSGPIYLIFECYCGDLLNLYRSKR 720
 Qy 708 EKFHRTWTEIFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDETE 767
 Db 721 EKFHRTWTEIFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDETE 780
 Qy 768 YENOKRLEEEEDLNVLTFEDLLCFAYQYAKGMEFLFEKSCVHRDLAARNVLTGHKVVKI 827
 Db 781 YENOKRLEEEEDLNVLTFEDLLCFAYQYAKGMEFLFEKSCVHRDLAARNVLTGHKVVKI 840
 Qy 828 CDFGLARDIMSDSNVYVRGNARLPVKWAPESLFGIYTIKSDVHSYGILLWEIFSLGVN 887
 Db 841 CDFGLARDIMSDSNVYVRGNARLPVKWAPESLFGIYTIKSDVHSYGILLWEIFSLGVN 900
 Qy 888 PYGIPVDANFYKLIQNGFKMDQPPYATEEIIYIMQSCWAFDSRRKRPFPNLTSLFGCOL 947
 Db 901 PYGIPVDANFYKLIQNGFKMDQPPYATEEIIYIMQSCWAFDSRRKRPFPNLTSLFGCOL 960
 Qy 948 ADAEAMYNQVNDGRVSECPHTYQNR 973
 Db 961 ADAEAMYNQVNDGRVSECPHTYQNR 986

RESULT 14

AAW63586

ID AAW63586 standard; Protein; 994 AA.

XX AC AAW63586;

XX DT 12-OCT-1998 (first entry)

XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:17.

XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
 XX KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.

OS Homo sapiens.
FN WO9817808-A1.
XX
XX 30-APR-1998.
XX
XX 13-OCT-1997; 97WO-JP03667.
XX
XX 18-OCT-1996; 96JP-0297329.
XX
XX (TAKI) TAKARA SHUZO CO LTD.
XX
XX Yokota S;
XX
XX WPI; 1998-362333/31.
DR N-PSDB; AAV39039.
XX
XX Nucleic acid sequences encoding receptor type protein kinase -
PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
PT M2
XX
XX Claim 6; Page 34-39; 80pp; Japanese.
XX
XX New nucleic acid sequences have been isolated which encode receptor type
CC protein kinases (especially a tyrosine kinases) having tandem repeats in
CC the juxtamembrane region. Also described in the present invention are:
CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)
CC representing FLT3 juxtamembrane receptor type protein kinases found in
CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences
CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
CC recognising the kinases or their portions including the tandem repeat
CC region; (3) nucleic acid sequences hybridising with the nucleic acids;
CC (4) a method for detecting the nucleic acid sequences in human tissue
CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
CC gene, and (iii) determining the size of the gene for comparing the size
CC of the normal gene not containing tandem repeats, and (5) kits for
CC carrying out the detection. The products and methods may be used for
CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.
XX
XX Sequence 994 AA:
XX
Query Match 97.7%; Score 5153.5; DB 19; Length 994;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 973; Conservative 0; Mismatches 0; Indels 21; Gaps 1;
QY 1 MPALARDAGTVP LLVVFSA MIFGTITNODLPV IKCVLINHKNDSSV GKSSSYPMVSESP 60
DB 1 MPALARDAGTVP LLVVFSA MIFGTITNODLPV IKCVLINHKNDSSV GKSSSYPMVSESP 60
QY 61 EDLGCALRQSSGTVY EAAAVEVDVSGASITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120
DB 61 EDLGCALRQSSGTVY EAAAVEVDVSGASITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120
QY 121 QNRGVVSMVLKMTETQAGYLLFIQSEATNYTLFTVSIKNTLLYTLRRPYFRKMNQD 180
DB 121 QNRGVVSMVLKMTETQAGYLLFIQSEATNYTLFTVSIKNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISSEVPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISSEVPEVVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNQTPQTLLPOLFLKVGEP LWRCKAVHVNHGFGLTWELNKALEBGNV FEM 300
DB 241 CTRFLTIDLNQTPQTLLPOLFLKVGEP LWRCKAVHVNHGFGLTWELNKALEBGNV FEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSSKHPQSQALVTIVGKGFINATNSSEDEY 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSSKHPQSQALVTIVGKGFINATNSSEDEY 360
QY 361 IDQYEEFCFSVRFKAYPQIRCTWTFTSRKSPCEOKGLDNGYSISKFCNKHKHPGEYIFHA 420
DB 361 IDQYEEFCFSVRFKAYPQIRCTWTFTSRKSPCEOKGLDNGYSISKFCNKHKHPGEYIFHA 420

QY 421 ENDDAQFTKMTFLNIRRRKQVLAESAASQSCFSDGYP LPSWTWKCKSDKSPNCTEEITE 480
DB 421 ENDDAQFTKMTFLNIRRRKQVLAESAASQSCFSDGYP LPSWTWKCKSDKSPNCTEEITE 480
QY 481 GYWNKANKRKVFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540
DB 481 GYWNKANKRKVFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540
QY 541 NISFYATIGVCLLFIWVLTLLICHYKKQFRYESQLQMVQVGTSSDNEYFYVDFREY --- 597
DB 541 NISFYATIGVCLLFIWVLTLLICHYKKQFRYESQLQMVQVGTSSDNEYFYVDFREY 600
QY 598 -----EYDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSI 639
DB 601 LKWEFPRENLEFGKVLGSEYDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSI 660
QY 640 QYAVKMLKEKADSSEREAALMSLKMTQLGSHENIVNLLGACTLSGPYILIFEYCCYGD 699
DB 661 QYAVKMLKEKADSSEREAALMSLKMTQLGSHENIVNLLGACTLSGPYILIFEYCCYGD 720
QY 700 LNYLSRKREKFRHTWTEIFKEHNFSEYPTFQSHPNSSMPGSRREVQIHPDSQISGLHGS 759
DB 721 LNYLSRKREKFRHTWTEIFKEHNFSEYPTFQSHPNSSMPGSRREVQIHPDSQISGLHGS 780
QY 760 FHSEDEIYENQKRLBEEDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLV 819
DB 781 FHSEDEIYENQKRLBEEDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLV 840
QY 820 THGKVYKICDFGLARDIMSDSNVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLW 879
DB 841 THGKVYKICDFGLARDIMSDSNVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLW 900
QY 880 EIFSLGVNPPGIPVDANFYKLIQNGFKMDQPFYATEEIIIMQSCWAFDSRKRPSFNL 939
DB 901 EIFSLGVNPPGIPVDANFYKLIQNGFKMDQPFYATEEIIIMQSCWAFDSRKRPSFNL 960
QY 940 TSFLGCOLADAEAMTQNVQDGRVSECPHTYQNR 973
DB 961 TSFLGCOLADAEAMTQNVQDGRVSECPHTYQNR 994
RESULT 15
AAR81868
ID AAR81868 standard; Protein: 1000 AA.
XX
AC AAR81868;
XX
DT 29-FEB-1996 (first entry)
XX
DE Flk2/flt3 tyrosine kinase receptor.
XX
KW Tyrosine kinase receptor: flk2; flt3; agonist; monoclonal antibody;
KW haematopoiesis; hypoplasemia; anaemia; thrombocytopenia; stem cell.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
XX Domain 1..542
XX FT /label= Extracellular_domain
XX
XX WO9527062-A1.
XX
XX PD 12-OCT-1995.
XX
XX PF 23-MAR-1995; 95WO-US03718.
XX
XX PR 04-APR-1994; 94US-0222299.
XX
XX PA (GETH) GENENTECH INC.
XX
XX PI Bennett BD, Broz SD, Matthews W, Zeigler FC;
XX
XX WPI; 1995-358636/46.
DR

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:28:59 ; Search time 24.0121 Seconds
(without alignments)
1216.760 Million cell updates/sec

Title: US-09-919-408-4
Perfect score: 5274
Sequence: 1 MPALARDAGRVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCFUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	1	US-07-977-451-4
2	5274	100.0	993	1	US-08-252-517-4
3	5274	100.0	993	1	US-07-906-397A-4
4	5274	100.0	993	1	US-08-601-891-4
5	5274	100.0	993	2	US-09-021-324-4
6	5274	100.0	993	5	PCT-US92-03893-4
7	5274	100.0	1160	5	PCT-US92-03401-4
8	5240	99.4	993	1	US-08-222-299-4
9	5240	99.4	993	2	US-08-434-878-4
10	5238	99.3	993	5	PCT-US95-03718-4
11	5238	99.3	993	1	US-08-183-211-2
12	5238	99.3	993	5	PCT-US95-00176A-2
13	4533.5	86.0	1000	1	US-08-222-299-2
14	4533.5	86.0	1000	2	US-08-434-878-2
15	4533.5	86.0	1000	5	PCT-US95-03718-2
16	4429.5	84.0	992	1	US-07-813-593-2
17	4429.5	84.0	992	1	US-07-977-451-2
18	4429.5	84.0	992	1	US-07-946-507-2
19	4429.5	84.0	992	1	US-08-252-517-2
20	4429.5	84.0	992	1	US-07-906-397A-2
21	4429.5	84.0	992	1	US-08-601-891-2
22	4429.5	84.0	992	2	US-09-021-324-2
23	4429.5	84.0	992	5	PCT-US92-02750-2
24	4429.5	84.0	992	5	PCT-US92-03401-2
25	4429.5	84.0	992	5	PCT-US92-03893-2
26	1654	31.4	481	4	US-07-912-122-4
27	1654	31.4	481	5	PCT-US93-06404-4

28 1251 23.7 976 3 US-08-750-141A-1 Sequence 1, Appli
29 1216.5 23.1 972. 3 US-08-750-141A-2 Sequence 2, Appli
30 1166 22.1 1089 1 US-08-180-195-36 Sequence 36, Appli
31 1166 22.1 1089 1 US-08-168-917-4 Sequence 4, Appli
32 1166 22.1 1089 1 US-08-477-323-36 Sequence 36, Appli
33 1166 22.1 1089 2 US-08-475-458-36 Sequence 36, Appli
34 1166 22.1 1089 2 US-08-460-510-4 Sequence 4, Appli
35 1166 22.1 1089 2 US-08-460-490-4 Sequence 4, Appli
36 1166 22.1 1089 3 US-08-980-400-36 Sequence 36, Appli
37 1166 22.1 1089 3 US-08-462-728-2 Sequence 2, Appli
38 1166 22.1 1089 4 US-09-583-458A-36 Sequence 36, Appli
39 1166 22.1 1089 4 US-09-583-210-36 Sequence 36, Appli
40 1166 22.1 1089 4 US-09-583-449A-36 Sequence 36, Appli
41 1166 22.1 1089 4 US-09-435-059-36 Sequence 36, Appli
42 1166 22.1 1089 4 US-08-461-917-2 Sequence 2, Appli
43 1166 22.1 1089 5 PCT-US92-00730-4 Sequence 4, Appli
44 1166 22.1 1089 5 PCT-US92-00862-4 Sequence 4, Appli
45 1086.5 20.6 1106 1 US-08-180-195-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-07-977-451-4
; Sequence 4, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-977-451-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVICKVLINHKNNDSVSGKSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVICKVLINHKNNDSVSGKSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120
Db 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120

Qy 121 QNRGVSMVILKMETQAGEVLLFIQSEATNTILFTVSIRNTLYTLRRPYFRKMEQND 180
Db 121 QNRGVSMVILKMETQAGEVLLFIQSEATNTILFTVSIRNTLYTLRRPYFRKMEQND 180

Qy 181 ALVCISEVPEPIVWLVLCDSOGESCKEESPAVWKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISEVPEPIVWLVLCDSOGESCKEESPAVWKEKVLHELFGTDIRCCARNELGRE 240

Qy 241 CTRLFTIDLNOTPQTLTLPOLFKVCEPLWIRCKAVVNHGFLGTLWELENKALEEGNYFEM 300
Db 241 CTRLFTIDLNOTPQTLTLPOLFKVCEPLWIRCKAVVNHGFLGTLWELENKALEEGNYFEM 300

Qy 301 STYSTNRTMIRLFAVSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360
Db 301 STYSTNRTMIRLFAVSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360

Qy 361 IDQVEEFCFSVRKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQGEYIIPA 420
Db 361 IDQVEEFCFSVRKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQGEYIIPA 420

Qy 421 ENDDAQFTKMTLNTIRRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNTIRRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480

Qy 481 GWNKRNKRVFGQWVSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIOD 540
Db 481 GWNKRNKRVFGQWVSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIOD 540

Qy 541 NISFYATIGVCLLFIWTLTLLCHIKYKQFRYESQLQMVQVVTGSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIWTLTLLCHIKYKQFRYESQLQMVQVVTGSSDNEYFYVDFREYED 600

Qy 601 LKWEPPRENLEFGKVLGSGAFQKVNATAYGISKTGVSITQAVKMLKEKADSSEREALMS 660
Db 601 LKWEPPRENLEFGKVLGSGAFQKVNATAYGISKTGVSITQAVKMLKEKADSSEREALMS 660

Qy 661 ELKMMTOLGSHENIYNLLGACTLSPVILFIPEYCCYGLDNLNLRKREKFRHTWTETIPE 720
Db 661 ELKMMTOLGSHENIYNLLGACTLSPVILFIPEYCCYGLDNLNLRKREKFRHTWTETIPE 720

Qy 721 HNFSYPTFQSHPNSSMPGSRVQTHPDSQISGLHGNFSHSEDETEYENQKRLBEEDL 780
Db 721 HNFSYPTFQSHPNSSMPGSRVQTHPDSQISGLHGNFSHSEDETEYENQKRLBEEDL 780

RESULT 2
US-08-252-517-4
; Sequence 4, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,517
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-252-517-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVKICVLINHKNDSSVGKSSYPWVSESP 60
DB 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVKICVLINHKNDSSVGKSSYPWVSESP 60
QY 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
DB 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
QY 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLYTLRPPYFRKMENQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLYTLRPPYFRKMENQD 180
QY 181 ALVCISESPEPIVWVLCDSQESCKEESPAVVKKEKYLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESPEPIVWVLCDSQESCKEESPAVVKKEKYLHELFGTDIRCCARNELGRE 240
QY 241 CTRLFIDLNQTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEEGNFEM 300
DB 241 CTRLFIDLNQTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEEGNFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSQALVTIVGKGFINATNSSDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSQALVTIVGKGFINATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYPQIRCTWFSRKSPPCOKGLDNGYSISKFCNHRKHQGEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPQIRCTWFSRKSPPCOKGLDNGYSISKFCNHRKHQGEYIFHA 420
QY 421 ENDDAOFKMTLNIRKPOVLAESAQSCFSDGVPPLSWTKKCSKSPNCTEITE 480
DB 421 ENDDAOFKMTLNIRKPOVLAESAQSCFSDGVPPLSWTKKCSKSPNCTEITE 480
QY 481 GVNRRKANRVFGQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GVNRRKANRVFGQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIVVLLILCHIKYKQPRYESQLOMQVQVTGSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIVVLLILCHIKYKQPRYESQLOMQVQVTGSSDNEYFYVDFREYED 600
QY 601 LKWEFFRENLEFGKVLGSGAFGVMNATAGISKTGVSIOAVKMLKEKADSSERREALMS 660
DB 601 LKWEFFRENLEFGKVLGSGAFGVMNATAGISKTGVSIOAVKMLKEKADSSERREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLFIEYCCYGDLLNLYLSRREKPHRTWTIEFKE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLFIEYCCYGDLLNLYLSRREKPHRTWTIEFKE 720
QY 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQDISGLHGNFSHSEDEIYEYENQKRLBEEEDL 780
DB 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQDISGLHGNFSHSEDEIYEYENQKRLBEEEDL 780
QY 781 NVLTFFEDLLCFAYQVAKGMEFFLFEKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMS 840
DB 781 NVLTFFEDLLCFAYQVAKGMEFFLFEKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYIGLLWEIFSLGPNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYIGLLWEIFSLGPNYPGIPVDANFYK 900

RESULT 3
US-07-906-397A-4
Sequence 4, Application US/07906397A
Patent No. 5621090
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,397A
FILING DATE: 19920626
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-906-397A-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVKICVLINHKNDSSVGKSSYPWVSESP 60
DB 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVKICVLINHKNDSSVGKSSYPWVSESP 60
QY 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
DB 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
QY 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLYTLRPPYFRKMENQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSIRNTLLYTLRPPYFRKMENQD 180
QY 181 ALVCISESPEPIVWVLCDSQESCKEESPAVVKKEKYLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESPEPIVWVLCDSQESCKEESPAVVKKEKYLHELFGTDIRCCARNELGRE 240
QY 241 CTRLFIDLNQTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEEGNFEM 300
DB 241 CTRLFIDLNQTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEEGNFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSQALVTIVGKGFINATNSSDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSQALVTIVGKGFINATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYPQIRCTWFSRKSPPCOKGLDNGYSISKFCNHRKHQGEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPQIRCTWFSRKSPPCOKGLDNGYSISKFCNHRKHQGEYIFHA 420
QY 421 ENDDAOFKMTLNIRKPOVLAESAQSCFSDGVPPLSWTKKCSKSPNCTEITE 480
DB 421 ENDDAOFKMTLNIRKPOVLAESAQSCFSDGVPPLSWTKKCSKSPNCTEITE 480
QY 481 GVNRRKANRVFGQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GVNRRKANRVFGQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIVVLLILCHIKYKQPRYESQLOMQVQVTGSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIVVLLILCHIKYKQPRYESQLOMQVQVTGSSDNEYFYVDFREYED 600
QY 601 LKWEFFRENLEFGKVLGSGAFGVMNATAGISKTGVSIOAVKMLKEKADSSERREALMS 660
DB 601 LKWEFFRENLEFGKVLGSGAFGVMNATAGISKTGVSIOAVKMLKEKADSSERREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLFIEYCCYGDLLNLYLSRREKPHRTWTIEFKE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLFIEYCCYGDLLNLYLSRREKPHRTWTIEFKE 720
QY 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQDISGLHGNFSHSEDEIYEYENQKRLBEEEDL 780
DB 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQDISGLHGNFSHSEDEIYEYENQKRLBEEEDL 780
QY 781 NVLTFFEDLLCFAYQVAKGMEFFLFEKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMS 840
DB 781 NVLTFFEDLLCFAYQVAKGMEFFLFEKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYIGLLWEIFSLGPNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYIGLLWEIFSLGPNYPGIPVDANFYK 900

61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIKNTLLYTLRPPYFRKMNQD 180
121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIKNTLLYTLRPPYFRKMNQD 180
181 ALVCISESPVPIVEMVLCDSOGESCKEESPAVVKKEKVLHFLFCTDIRCCARNELGRE 240
181 ALVCISESPVPIVEMVLCDSOGESCKEESPAVVKKEKVLHFLFCTDIRCCARNELGRE 240
241 CTRLEFDLNTQPTTLPQLFLKVGEPWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
241 CTRLEFDLNTQPTTLPQLFLKVGEPWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
301 STYSTNRTMIRILFAPVSSVARNDTGYTSCSKHPSQSALVTIVCKGINATNSSEYD 360
301 STYSTNRTMIRILFAPVSSVARNDTGYTSCSKHPSQSALVTIVCKGINATNSSEYD 360
361 IDOYEFCFSVRKAYPQIRCTWTFSRKSPCEQKLDNGYSISKFCNKHQGEYIFHA 420
361 IDOYEFCFSVRKAYPQIRCTWTFSRKSPCEQKLDNGYSISKFCNKHQGEYIFHA 420
421 ENDDAQFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
421 ENDDAQFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
481 GVNRRKANRKGQVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGFFFIQD 540
481 GVNRRKANRKGQVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGFFFIQD 540
541 NISFYATIGVCLLFIVVLTLLCHKVKKQFRYESQVMQVTSQSSDNEYFYDFREYD 600
541 NISFYATIGVCLLFIVVLTLLCHKVKKQFRYESQVMQVTSQSSDNEYFYDFREYD 600
601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIOQVAVKMLKEKADSSEREA 660
601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIOQVAVKMLKEKADSSEREA 660
661 ELKMTQLGSHENIVNLLGACTLSGPIYLLIFECYCGDLLNLYLSRREKPHRTWTEIFKE 720
661 ELKMTQLGSHENIVNLLGACTLSGPIYLLIFECYCGDLLNLYLSRREKPHRTWTEIFKE 720
721 HNFSPYPTQSHNSMSPGSGREVQIHPDSQISGLHGNPSHSEDEIENOKRLEEDL 780
721 HNFSPYPTQSHNSMSPGSGREVQIHPDSQISGLHGNPSHSEDEIENOKRLEEDL 780
781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
841 NYVVRGNARLPVKWMAPESLFEGITTKSDVSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
841 NYVVRGNARLPVKWMAPESLFEGITTKSDVSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
901 LIQNGFKMDQPFVATEIYIIMSCWAFDSCRKPSFNLTSFLGCQLADAEAEAMYQNVGD 960
901 LIQNGFKMDQPFVATEIYIIMSCWAFDSCRKPSFNLTSFLGCQLADAEAEAMYQNVGD 960
961 RVSECPHTYQNRFPFSEMDLGLLSPQAQVEDS 993
961 RVSECPHTYQNRFPFSEMDLGLLSPQAQVEDS 993

RESULT 4

US-08-601-891-4
: Sequence 4, Application US/08601891
: Patent No. 5747651
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/601,891
FILING DATE: 15-FEB-1996
CLASSIFICATION: 530
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 13-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-601-891-4

Query Match 100.0%; Score 5274; DB 1; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAIMFGTITNQDLPVIRKVLINHNKNDSSVKGSSYPMVSESP 60
DB 1 MPALARDAGTVPLLVVFSAIMFGTITNQDLPVIRKVLINHNKNDSSVKGSSYPMVSESP 60
QY 61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
DB 61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
QY 121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIKNTLLYTLRPPYFRKMNQD 180
DB 121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIKNTLLYTLRPPYFRKMNQD 180

181 ALVCISESPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
181 ALVCISESPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
241 CTRFTIDLNQOTPTTLPOLFLKVGEPWLRCKAVHNVHGFGLTWELKALEEGNFEM 300
241 CTRFTIDLNQOTPTTLPOLFLKVGEPWLRCKAVHNVHGFGLTWELKALEEGNFEM 300
301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSOSALVTIVGKGFNATNSSDEYE 360
301 STYSTNRTMIRILFAFVSSVARNDGTGYTSCSSKHPQSOSALVTIVGKGFNATNSSDEYE 360
361 IDOYEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNKHKHPGCEYIFHA 420
361 IDOYEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNKHKHPGCEYIFHA 420
421 ENDDAOFKFTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
421 ENDDAOFKFTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480
481 GVNRRKANRVFGOWSSSTLNKSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIOD 540
481 GVNRRKANRVFGOWSSSTLNKSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIOD 540
541 NISFYATIGVCLLFIVVLTLLCHIKYKQFRESQLOMVQVGTSSDNEYFYVDFREYEYD 600
541 NISFYATIGVCLLFIVVLTLLCHIKYKQFRESQLOMVQVGTSSDNEYFYVDFREYEYD 600
601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTVSIOVAVKMLKEKADSSREALMS 660
601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTVSIOVAVKMLKEKADSSREALMS 660
661 ELKMTOLGSHENIVNLGACTLSGPIYLIFECYCGDILLNLRKSKREKPHRTWTEIFKE 720
661 ELKMTOLGSHENIVNLGACTLSGPIYLIFECYCGDILLNLRKSKREKPHRTWTEIFKE 720
721 HNFSEFYPTFQSHNSMPGSRVQIHPDSQISGLHGNFSHSEDEIEYENOKRLEEDL 780
721 HNFSEFYPTFQSHNSMPGSRVQIHPDSQISGLHGNFSHSEDEIEYENOKRLEEDL 780
781 NVLTFEDLLCFAYOVAKGMFLEBKSCVHRDLAARNVLTGHKVVKICDGLARDIMSDS 840
781 NVLTFEDLLCFAYOVAKGMFLEBKSCVHRDLAARNVLTGHKVVKICDGLARDIMSDS 840
841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGLLWEIFSLGVNPNYPGIPVDANFYK 900
841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGLLWEIFSLGVNPNYPGIPVDANFYK 900
901 LIONGFKMDOPFYATBEIYIIMOSWAFDSKRKPSFNLTSFLGCCOLADAEAMQVNDG 960
901 LIONGFKMDOPFYATBEIYIIMOSWAFDSKRKPSFNLTSFLGCCOLADAEAMQVNDG 960
961 RVSECPHTYQNRPPFRSREMDLGLLSQAVQEDS 993
961 RVSECPHTYQNRPPFRSREMDLGLLSQAVQEDS 993

RESULT 5

US-09-021-324-4
; Sequence 4, Application US/09021324
; Patent No. 5912133
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Thor R.
; TITLE OF INVENTION: TORIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021.324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/977.451
FILING DATE: 1992-11-19
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906.397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813.593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793.065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728.913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679.666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-324-4

Query Match 100.0%; Score 5274; DB 2; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVWFSAIFGTITNODLPVVKCVLINHKNDSSVKGSSSPVWSESP 60
Db 1 MPALARDAGTVPLLVWFSAIFGTITNODLPVVKCVLINHKNDSSVKGSSSPVWSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLOVLVDAPGNISCLWFKHSSLNCOHPFDL 120
Db 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLOVLVDAPGNISCLWFKHSSLNCOHPFDL 120
QY 121 QNRGVWSMVLKMTQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
Db 121 QNRGVWSMVLKMTQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISESPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
Db 181 ALVCISESPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
QY 241 CTRFTIDLNQOTPTTLPOLFLKVGEPWLRCKAVHNVHGFGLTWELKALEEGNFEM 300
Db 241 CTRFTIDLNQOTPTTLPOLFLKVGEPWLRCKAVHNVHGFGLTWELKALEEGNFEM 300

Db 241 CTRLETTDLNQTPQTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCOKGLDNGYSISKFCNKHQPGGEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCOKGLDNGYSISKFCNKHQPGGEYIFHA 420
QY 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCSKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCSKSPNCTEITE 480
QY 481 GVMNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GVMNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLICHKYKQFRIESQLOMQVQVTSNDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLICHKYKQFRIESQLOMQVQVTSNDNEYFYVDFREYED 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
Db 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
QY 661 ELKMWTLQSHENIVNLLGACTLSGPIYLIFEYCCYDOLLNYSKREKFRHTWTEIFE 720
Db 661 ELKMWTLQSHENIVNLLGACTLSGPIYLIFEYCCYDOLLNYSKREKFRHTWTEIFE 720
QY 721 HNFsfyptfQSHPNSSMPGSRVQIHPDSQOISGLHGNFSHSEDEIYENOKRLEEDL 780
Db 721 HNFsfyptfQSHPNSSMPGSRVQIHPDSQOISGLHGNFSHSEDEIYENOKRLEEDL 780
QY 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKVTKCDFGLARDIMSDS 840
Db 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKVTKCDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYIGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYIGILLWEIFSLGVNPNYPGIPVDANFYK 900
QY 901 LIQNGFKMDOPFYATIEIYIMOSCAFDSKRPSPNLTSLFGLCCOLADAEAMYNQVVG 960
Db 901 LIQNGFKMDOPFYATIEIYIMOSCAFDSKRPSPNLTSLFGLCCOLADAEAMYNQVVG 960
QY 961 RVSECPHTYQNRPRFSGREMDGLLSPOAQVEDS 993
Db 961 RVSECPHTYQNRPRFSGREMDGLLSPOAQVEDS 993

RESULT 6
PCT-US92-09893-4
; Application PC/TUS9209893
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TOTIPORENT HEMATOPOIETIC STEM CELL
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09893

FILING DATE: 19921116
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7PT
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09893-4
Query Match 100.0%; Score 5274; DB 5; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVLINHKNDSSVSKSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVLINHKNDSSVSKSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
Db 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
QY 121 QNRGVVSMWILKMTQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMEQD 180
Db 121 QNRGVVSMWILKMTQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMEQD 180
QY 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCARNELGRE 240
Db 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCARNELGRE 240
QY 241 CTRLETTDLNQTPQTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
Db 241 CTRLETTDLNQTPQTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCOKGLDNGYSISKFCNKHQPGGEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCOKGLDNGYSISKFCNKHQPGGEYIFHA 420
QY 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCSKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCSKSPNCTEITE 480
QY 481 GVMNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GVMNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLICHKYKQFRIESQLOMQVQVTSNDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLICHKYKQFRIESQLOMQVQVTSNDNEYFYVDFREYED 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
Db 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
QY 661 ELKMWTLQSHENIVNLLGACTLSGPIYLIFEYCCYDOLLNYSKREKFRHTWTEIFE 720
Db 661 ELKMWTLQSHENIVNLLGACTLSGPIYLIFEYCCYDOLLNYSKREKFRHTWTEIFE 720
QY 721 HNFsfyptfQSHPNSSMPGSRVQIHPDSQOISGLHGNFSHSEDEIYENOKRLEEDL 780
Db 721 HNFsfyptfQSHPNSSMPGSRVQIHPDSQOISGLHGNFSHSEDEIYENOKRLEEDL 780

QY 781 NVLTFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840
QY 841 NYVVGARLPVKWAPESLFEIYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVGARLPVKWAPESLFEIYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGKMDOPFYATEEIIIMQSCWAFDSKRPSPNLTSLFGLCQLADAEAMYQNVDG 960
DB 901 LIQNGKMDOPFYATEEIIIMQSCWAFDSKRPSPNLTSLFGLCQLADAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 7

PC1-US92-05401-4
; Sequence 4, Application PC/TUS9205401
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05401
; FILING DATE: 19920626

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1160 amino acids
; TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

PC1-US92-05401-4

Query Match 100.0%; Score 5274; DB 5; Length 1160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVCLINHKNDSSVGKSSSYPMVSESP 60
DB 20 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVCLINHKNDSSVGKSSSYPMVSESP 79
QY 61 EDLGCALRQSSGTVYAAAVEVDVSASITLQVLVDAPGNISCLWFKHSSLCNQPHFDL 120
DB 80 EDLGCALRQSSGTVYAAAVEVDVSASITLQVLVDAPGNISCLWFKHSSLCNQPHFDL 139
QY 121 QNRGVVSMVLKMTQAGEYLLFIQSEATNYTLFTVSRINLTLLRPPYFRKMNQD 180
DB 140 QNRGVVSMVLKMTQAGEYLLFIQSEATNYTLFTVSRINLTLLRPPYFRKMNQD 199
QY 181 ALVCISESVPEPIVWVCLDSQGESCKEESPAVVVKEEVLHFGTDIRCCARNELGRE 240

DB 200 ALVCISESVPEPIVWVCLDSQGESCKEESPAVVVKEEVLHFGTDIRCCARNELGRE 259
QY 241 CTRUFTIDLNPOTTLPLQLFLKVGEPLEWIRCKAVVHVHGFGLTWELNKALEEGNYFDM 300
DB 260 CTRUFTIDLNPOTTLPLQLFLKVGEPLEWIRCKAVVHVHGFGLTWELNKALEEGNYFDM 319
QY 301 STYSTNRMTIRILFAFVSVVARNDTGYTCSKSPKPSOSALVTIVGKGFINATNSEDYE 360
DB 320 STYSTNRMTIRILFAFVSVVARNDTGYTCSKSPKPSOSALVTIVGKGFINATNSEDYE 379
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKGHPGYIFHA 420
DB 380 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKGHPGYIFHA 439
QY 421 ENDDAQFTKMTFLAIRRKPQVLAELASASQSCFSDGYPLPSPWTKKCSKSPNCTEETE 480
DB 440 ENDDAQFTKMTFLAIRRKPQVLAELASASQSCFSDGYPLPSPWTKKCSKSPNCTEETE 499
QY 481 GWNRKANKRVFGOWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
DB 500 GWNRKANKRVFGOWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 559
QY 541 NISFYATIGVCLLFIVVLTLLICHKKQFRYESQLOVQVQVQVQVQVQVQVQVQVQVQV 600
DB 560 NISFYATIGVCLLFIVVLTLLICHKKQFRYESQLOVQVQVQVQVQVQVQVQVQVQV 619
QY 601 LKWEPPRENLEFGKVLGSGAFKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 660
DB 620 LKWEPPRENLEFGKVLGSGAFKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 679
QY 661 ELKWMTOQLGSHENIVNLLGACTLSGPIYLIPEYCYGDLNLYLRSKRKFHRTWTEIFKE 720
DB 680 ELKWMTOQLGSHENIVNLLGACTLSGPIYLIPEYCYGDLNLYLRSKRKFHRTWTEIFKE 739
QY 721 HNFIFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHNSFSEDEIEYENOKRLEEDL 780
DB 740 HNFIFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHNSFSEDEIEYENOKRLEEDL 799
QY 781 NVLTFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840
DB 800 NVLTFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 859
QY 841 NYVVGARLPVKWAPESLFEIYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 860 NYVVGARLPVKWAPESLFEIYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 919
QY 901 LIQNGKMDOPFYATEEIIIMQSCWAFDSKRPSPNLTSLFGLCQLADAEAMYQNVDG 960
DB 920 LIQNGKMDOPFYATEEIIIMQSCWAFDSKRPSPNLTSLFGLCQLADAEAMYQNVDG 979
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 980 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 1012

RESULT 8

US-08-222-299-4
; Sequence 4, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:

APPLICANT: Bennett, Brian D.
APPLICANT: Broz, Susan D.
APPLICANT: Matthews, William
APPLICANT: Zeigler, Francis C.
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,299
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-299-4

Query Match 99.4%; Score 5240; DB 1; Length 993;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVFSAIFGTITNQDLPVTKVNLNHNNDSSVKGSSYPMVSESP 60
DB 1 MPALARDGGQLPLLVFSAIFGTITNQDLPVTKVNLNHNNDSSVKGSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVYAAAVEVDVSAITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVYAAAVEVDVSAITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
QY 121 QNRGVSMWILKMTQTQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVSMWILKMTQTQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180
QY 181 ALVCISESVPEPIVWVLCDSQSGESCKEESPAVVVKKEVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQSGESCKEESPAVVVKKEVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLTIDLTQTPOTTLPOLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
DB 241 CTRLTIDLTQTPOTTLPOLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCCSSKHPSQSALVTIVGKFINATNSSDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTCCSSKHPSQSALVTIVGKFINATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYQIRCTWTFSRKSFCQKGLDNGYSISKFCNKHKHPGEYIFHA 420
DB 361 IDQYEEFCFSVRKAYQIRCTWTFSRKSFCQKGLDNGYSISKFCNKHKHPGEYIFHA 420
QY 421 ENDDAQFTKMTLIRKRPQVLAESAQSCFSDGVPPLSWTWKCKSDKSPNCTEEITE 480
DB 421 ENDDAQFTKMTLIRKRPQVLAESAQSCFSDGVPPLSWTWKCKSDKSPNCTEEITE 480
QY 481 GVMNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLTGTCETILLNSPGFPFFIQD 540
DB 481 GVMNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLTGTCETILLNSPGFPFFIQD 540
QY 541 NISFYATIGVCLLFIIVTLILCHIKYKQFRYESQLQWQVGTSSDNEYFYVDREYED 600
DB 541 NISFYATIGVCLLFIIVTLILCHIKYKQFRYESQLQWQVGTSSDNEYFYVDREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLKEKADSSREALMS 660

DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLKEKADSSREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEXCCYGDLLNLYLRKREKFRHTWTEIFKE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEXCCYGDLLNLYLRKREKFRHTWTEIFKE 720
QY 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQISGLHGNSFHSDEIEYENQKRLEEEEDL 780
DB 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQISGLHGNSFHSDEIEYENQKRLEEEEDL 780
QY 781 NVLTFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPELSLEFEGIYTIKSDVMSYGIILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPELSLEFEGIYTIKSDVMSYGIILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPFYATEEIIYIMQSWAFDSKRKPSFPNLTSLFLGCQLADAEAMYNQVNDG 960
DB 901 LIQNGFKMDQPFYATEEIIYIMQSWAFDSKRKPSFPNLTSLFLGCQLADAEAMYNQVNDG 960
QY 961 RVSECPHTYQNRPRFSDREMDLGLLSPQAQVEDS 993
DB 961 PVSECPHTYQNRPRFSDREMDLGLLSPQAQVEDS 993

RESULT 9
US-08-434-878-4
Sequence 4, Application US/08434878
Patent No. 5997865
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Broz, Susan D.
APPLICANT: Matthews, William
APPLICANT: Zeigler, Francis C.
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,878
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-434-878-4

Query Match 99.4%; Score 5240; DB 2; Length 993;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNQDLPVIRKVLINHKNDSSVCKSSYPVWSESP 60
Db 1 MPALARDGGQLPLLVFSAMIFGTITNQDLPVIRKVLINHKNDSSVCKSSYPVWSESP 60

QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
Db 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

QY 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180
Db 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVYKKEKVLHFGMDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVYKKEKVLHFGMDIRCCARNELGRE 240

QY 241 CTRFLTIDLNTQPTTLPOLFLKVGEPILWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300
Db 241 CTRFLTIDLNTQPTTLPOLFLKVGEPILWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360

QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420

QY 421 ENDDAQFTKMTILNIRKPOVLAESASQSCFSDGYPLEPSWTKKCDSPNCTEITE 480
Db 421 ENDDAQFTKMTILNIRKPOVLAESASQSCFSDGYPLEPSWTKKCDSPNCTEITE 480

QY 481 GVNRRKANRKFQWVSSSTLNKSEATKGLVKCCAYNSLGTSCETILLNSPGPEPIQD 540
Db 481 GVNRRKANRKFQWVSSSTLNKSEATKGLVKCCAYNSLGTSCETILLNSPGPEPIQD 540

QY 541 NTSFYATIGVCLLFIIVLTLILCHYKKQRYESQLOMQVQVGTSSDNEYFYVDFREYED 600
Db 541 NTSFYATIGVCLLFIIVLTLILCHYKKQRYESQLOMQVQVGTSSDNEYFYVDFREYED 600

QY 601 LKWEPPRENLEFKVLGSAFGKVMNATAYGISTGYSIQVAVKMLKEKADSSEREALMS 660
Db 601 LKWEPPRENLEFKVLGSAFGKVMNATAYGISTGYSIQVAVKMLKEKADSSEREALMS 660

QY 661 ELKMMTQLGSHENIVNLGACTLSGPYLLIFCYCCYDGLNLYRSRKEKPHRTWTTFKE 720
Db 661 ELKMMTQLGSHENIVNLGACTLSGPYLLIFCYCCYDGLNLYRSRKEKPHRTWTTFKE 720

QY 721 HNFSFYPTQSHPNSSMPGSRREVQIHPDSQIISGLHGNPSHSEDEIYENQKLEEBEDL 780
Db 721 HNFSFYPTQSHPNSSMPGSRREVQIHPDSQIISGLHGNPSHSEDEIYENQKLEEBEDL 780

QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDPLGARDIMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDPLGARDIMSDS 840

QY 841 NVVVRGNARLPVKWMAPESLFEGIYITKSDWSYIGILLWEIFSLGVNYPYGPVNDANFYK 900
Db 841 NVVVRGNARLPVKWMAPESLFEGIYITKSDWSYIGILLWEIFSLGVNYPYGPVNDANFYK 900

QY 901 LIQNGFKMDQPFYATBEIYIIMOSWAFDSRKRPSFNLTSFLGCQLADAEAEAMYQNVDG 960
Db 901 LIQNGFKMDQPFYATBEIYIIMOSWAFDSRKRPSFNLTSFLGCQLADAEAEAMYQNVDG 960

QY 961 RVSECPHTQNRPPFSREMDLGLLSFOAQVEDS 993
Db 961 PVSECPHTYQNRPPFSREMDLGLLSFOAQVEDS 993

RESULT 10
PCT-US95-03718-4
; Sequence 4: Application PC/TUS9503718
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 879PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US95-03718-4

Query Match 99.4%; Score 5240; DB 5; Length 993;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNQDLPVIRKVLINHKNDSSVCKSSYPVWSESP 60
Db 1 MPALARDGGQLPLLVFSAMIFGTITNQDLPVIRKVLINHKNDSSVCKSSYPVWSESP 60

QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
Db 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

QY 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180
Db 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVYKKEKVLHFGMDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVYKKEKVLHFGMDIRCCARNELGRE 240

QY 241 CTRFLTIDLNTQPTTLPOLFLKVGEPILWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300
Db 241 CTRFLTIDLNTQPTTLPOLFLKVGEPILWIRCKAVHNHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360

QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420

Qy 421 ENDDAQFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480
Db 421 ENDDAQFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480
Qy 481 GYMNKRANKVFGQWSSSTLNNSEAIKGFVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Db 481 GYMNKRANKVFGQWSSSTLNNSEAIKGFVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLCHYKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLCHYKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 600
Qy 601 LKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREAALMS 660
Db 601 LKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREAALMS 660
Qy 661 ELKMTQGLSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYLRSKREKPHRTWTEIFKE 720
Db 661 ELKMTQGLSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYLRSKREKPHRTWTEIFKE 720
Qy 721 HNFSEFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSSEDEIYEYENOKRLEED 780
Db 721 HNFSEFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSSEDEIYEYENOKRLEED 780
Qy 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 840
Qy 841 NYVVRGNARLPVKWMAPELSLFEIGIYITKSDVSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPELSLFEIGIYITKSDVSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Qy 901 LIQNGFKMDOPFYATEIYIIMOSCAFDSRKRPPNLTSLFLGCOLADAEAEAMYNQVNDG 960
Db 901 LIQNGFKMDOPFYATEIYIIMOSCAFDSRKRPPNLTSLFLGCOLADAEAEAMYNQVNDG 960
Qy 961 RVSECPHTYONRRPFRSREMDLGLSLPQAQVEDS 993
Db 961 PVSECPHTYONRRPFRSREMDLGLSLPQAQVEDS 993

RESULT 11
US-08-183-211-2
Sequence 2, Application US/08183211
Patent No. 5618709
GENERAL INFORMATION:
APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR
TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STK-1 PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
STREET: Suite 1800, Penn Center Plaza
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/183,211
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 3957-15
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5618709e
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-183-211-2

Query Match 99.3%; Score 5238; DB 1; Length 993;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60
Qy 61 EDLGCALRPOSSCTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHDL 120
Db 61 EDLGCALRPOSSCTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHDL 120
Qy 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSTRNTLLYTLRPYFRKMNQD 180
Db 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSTRNTLLYTLRPYFRKMNQD 180
Qy 181 ALVCISESPEPVEVWVLCDSQGESCKEESPAVVVKEEVLHFGTDIRCCARNELGRE 240
Db 181 ALVCISESPEPVEVWVLCDSQGESCKEESPAVVVKEEVLHFGTDIRCCARNELGRE 240
Qy 241 CTRLTIDLTNOTPTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEGNFEM 300
Db 241 CTRLTIDLTNOTPTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEGNFEM 300
Qy 301 STYSTNRTMIRILFAVSSVARNDTGYTCTSSSKHPQSALVTIVGKFINATNSSDEYE 360
Db 301 STYSTNRTMIRILFAVSSVARNDTGYTCTSSSKHPQSALVTIVGKFINATNSSDEYE 360
Qy 361 IDOYEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKHOPEYIFH- 419
Db 361 IDOYEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKHOPEYIFH- 420
Qy 420 AENDDAQFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEIT 479
Db 420 AENDDAQFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEIT 479
Qy 480 EGVNKRANKVFGQWSSSTLNNSEAIKGFVKKCCAYNSLGTSCETILLNSPGPPFIQ 539
Db 480 EGVNKRANKVFGQWSSSTLNNSEAIKGFVKKCCAYNSLGTSCETILLNSPGPPFIQ 539
Qy 540 DNISFYATIGVCLLFIVVLTLLCHYKQFRYESQLOMVQVGTSSDNEYFYVDFREY 599
Db 540 DNISFYATIGVCLLFIVVLTLLCHYKQFRYESQLOMVQVGTSSDNEYFYVDFREY 599
Qy 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREAALM 659
Db 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREAALM 659
Qy 660 SELKMTQGLSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYLRSKREKPHRTWTEIFK 719
Db 660 SELKMTQGLSHENIVNLLGACTLSGPIYILFEYCCYGDLLNLYLRSKREKPHRTWTEIFK 719
Qy 720 EHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSSEDEIYEYENOKRLEED 779
Db 720 EHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSSEDEIYEYENOKRLEED 779
Qy 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 839
Db 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 839
Qy 840 SNYVVRGNARLPVKWMAPELSLFEIGIYITKSDVSYGILLWEIFSLGVNPNYPGIPVDANFY 899

Db 840 SNVVRGNARLPVKMAPESLFEGYTIKSDVWSYGILLWEIFSLGVNPGIPVDANFY 899
QY 900 KLIONGFKMDOPPYATEEYIIIMQSCWAFDSRKRPSPNLTSLGCOLADAEAMQONVD 959
Db 900 KLIONGFKMDOPPYATEEYIIIMQSCWAFDSRKRPSPNLTSLGCOLADAEAMQONVD 959
QY 960 GRVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993
Db 960 GRVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 12
PCT-US95-00176A-2
; Sequence 2, Application PC/TUS9500176A
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR INHIBITING EXPRESSION OF THE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00176A
; FILING DATE: 6 January 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,211
; FILING DATE: 14 January 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-14 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
PCT-US95-00176A-2

Query Match 99.3%; Score 5238; DB 5; Length 993;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVWSESP 60
Db 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVWSESP 60
QY 61 EDLGCALRPQSSGTVYEAADVEVDVSAITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
Db 61 EDLGCALRPQSSGTVYEAADVEVDVSAITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
QY 121 QNRGVVSMVILKMTQTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180
Db 121 QNRGVVSMVILKMTQTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVYKKEKVLHELFGTDIRCCARNELGRE 240

Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVYKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLFTTDLNQTPTTLPOLFLKVGEPDLWIRKAVVHNHGGFLTWELNKALEBGNFYEM 300
Db 241 CTRLFTTDLNQTPTTLPOLFLKVGEPDLWIRKAVVHNHGGFLTWELNKALEBGNFYEM 300
QY 301 STYSTNRTMIRILFAFYSSVARNDTGYTTCSSSKHPQSALVTVIGKGFINATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFYSSVARNDTGYTTCSSSKHPQSALVTVIGKGFINATNSSEDEY 360
QY 361 IDOYEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHPQGEYIEHA 420
Db 361 IDOYEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHPQGEYIEHA 420
QY 420 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDSPNCTEIT 479
Db 421 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDSPNCTEIT 479
QY 480 EGVNRRKANRVFGQWYSSSTLNNSAIGKFLVKCCAYNSLGTSCETILLNSPGPFPIQ 539
Db 480 EGVNRRKANRVFGQWYSSSTLNNSAIGKFLVKCCAYNSLGTSCETILLNSPGPFPIQ 539
QY 540 DNISFYATIGVCLLFIVVLTLLICHYKQFRYESQSQMVGVTGSSDNEYFYVDFREY 599
Db 540 DNISFYATIGVCLLFIVVLTLLICHYKQFRYESQSQMVGVTGSSDNEYFYVDFREY 599
QY 600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYIGSKTGVSIQVAVKMLKEKADSSREALM 659
Db 600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYIGSKTGVSIQVAVKMLKEKADSSREALM 659
QY 660 SELKMMTQLGSHENIVNLGACTLSGPTYLIFCYCCYGDLLNLYRSKREKFRHTWTEIFK 719
Db 660 SELKMMTQLGSHENIVNLGACTLSGPTYLIFCYCCYGDLLNLYRSKREKFRHTWTEIFK 719
QY 720 EHNFSFYPTQSHPSNMPGSRVQIHPDSDQISGLHGNSEFHSDEIEYENQKLEED 779
Db 720 EHNFSFYPTQSHPSNMPGSRVQIHPDSDQISGLHGNSEFHSDEIEYENQKLEED 779
QY 780 LNVLTFFEDLLCFAYQVAKGMEFFKSCVHRDLAARNVLYTHGKVKVTCDFGLARDINSD 839
Db 780 LNVLTFFEDLLCFAYQVAKGMEFFKSCVHRDLAARNVLYTHGKVKVTCDFGLARDINSD 839
QY 840 SNYVVRGNARLPVKMAPESLFEGYTIKSDVWSYGILLWEIFSLGVNPGIPVDANFY 899
Db 840 SNYVVRGNARLPVKMAPESLFEGYTIKSDVWSYGILLWEIFSLGVNPGIPVDANFY 899
QY 900 KLIONGFKMDOPPYATEEYIIIMQSCWAFDSRKRPSPNLTSLGCOLADAEAMQONVD 959
Db 900 KLIONGFKMDOPPYATEEYIIIMQSCWAFDSRKRPSPNLTSLGCOLADAEAMQONVD 959
QY 960 GRVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993
Db 960 GRVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 13
US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,299
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1000 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-299-2

Query Match 86.0%; Score 4533.5; DB 1; Length 1000;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVPLLVWFAMFGTITNODLPVVKVNLNHNKNDSSVGSYPVWVS 59
DB 1 MPALAQRSDRRLLLVLSVMTLETITNODLPVVKVNLNHNKNDSSVGSYPVWVS 60
QY 60 PEDLCALRQSSGTYEAAVEVDVSASITLOVLVDAPGNISCLWVFKHSSLCOPHED 119
DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLOVLATPGDLSCLWVFKHSSLCOPHED 120
QY 120 LQNRGVSVMLKMTQAGEYLLFTOSBATNTILFTVSRINTLYTLRRPYFRKMENQ 179
DB 121 LQNRGVSMALLNVTQAGEYLLHTIOSEAAANYTLFTVNRDTQLYLRRPYFRKMENQ 180
QY 180 DALVCISESVPEPIVEMVLCDSGESCCKEESPAVVKKEKVLHFGTDTRCCARNELGR 239
DB 181 DALLCISEGPEPTVEWVLCSSHRESCKEEGPAVVKKEKVLHFGTDTRCCARNALGR 240
QY 240 ECTRLFTIDNTPOTTLPOLFLKVGEPWIRCKAVHVNHGFLTWELNKALEEGNYFE 299
DB 241 ESTKFTIDLNQAPSTLPOLFLKVGEPWIRCKAVHVNHGFLTWELNKALEEGSYFE 300
QY 300 MSTYSTNRTMIRLLFAFVSSVARNDTGYTCSSSKHPSQSALVTIVKGFINATNSSEDY 359
DB 301 MSTYSTNRTMIRLLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSOEY 360
QY 360 EIDYEEFCFSVRFKAYPOLRCTWTFSRSPCEKGLDNGYSISKPCNHKHPGGEYIFH 419
DB 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQFCEQRLDGYISKPCDHNKHPGEYIFY 420
QY 420 AENDDAQFTKMTNIRRKPOVLAEASASQSCFSDGYPLPSTWTKCKSDKSPNCTEIT 479
DB 421 AENDDAQFTKMTNIRKPOVLANKASASQSCSDGYPLPSTWTKCKSDKSPNCTEIP 480
QY 480 EGYNRKANRKFQGWVSSSTLNSEAKIFGLVKCCAYNSLGTSCETILLNSGPPFFFIQ 539
DB 481 EGYVNNKANRKFQGWVSSSTLNSEAGKLLVKCCAYNSMGTSCTIFLNSGPPFFFIQ 540
QY 540 DNISFATIGVCLLFIIVLTLCHYKKQFRYESQLOMQVQTCSSDNEYFYVDREY 599
DB 541 DNISFATIGVCLLFIIVLTLCHYKKQFRYESQLOMQVQTCSSDNEYFYVDREY 600
QY 600 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIQVAVKMLKEKADSCKEALM 659

DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIQVAVKMLKEKADSCKEALM 660
QY 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKRKFHRTWTEIFK 719
DB 661 SELKMTHLGHDNIVNLLGACTLSGPIYLIFEYCCYDGLLNLRSKRKFHRTWTEIFK 720
QY 720 EHNFSYPTFQSHPNSSMPGSRREVQIHPSDQISGLHNGSFHSEDEIEYENQKRL--EEE 777
DB 721 EHNFSYPTFQSHPNSSMPGSRREVQIHPSDQISGLHNGSFHSEDEIEYENQKRLAEEE 780
QY 778 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIM 837
DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
QY 838 SDSNVVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGLLWEIFSLGVNPNYPIGPVDAN 897
DB 841 SDSNVVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGLLWEIFSLGVNPNYPIGPVDAN 900
QY 898 FYKLIONGFKMDOPFYATEEIIYIMOSWAFDSKRKRPFPNLTSLFCCOLADAEEMAYQN 957
DB 901 FYKLIOGFKMEOPFYATEEIIYIMOSWAFDSKRKRPFPNLTSLFCCOLADAEEMAYQN 960
QY 958 VQGRVSECHTYONRRPFRSEMDLGLLSPOAOYE 991
DB 961 MCGNVPEHPSIYQNRRLPSREAGSEPPSPQAOVK 994

RESULT 14
US-08-434-878-2
Sequence 2, Application US/08434878
Patent No. 5997865
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Broz, Susan D.
APPLICANT: Matthews, William
APPLICANT: Zeigler, Francis C.
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,878
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1000 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-434-878-2

Query Match 86.0%; Score 4533.5; DB 2; Length 1000;

Best Local Similarity 85.6%; Pred. No. 0;
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVPLLVVFSAMIFGTITNODLPVVKCVLINHKNNDSSVGSXSSYPVMS 59
Db 1 MRALAQRSDRRLLLVVLSVMIETVNTODLPVVKCVLINSHENNGSSAGKPSRYMVRGS 60
QY 60 PEDLGCALRQSSGTVYEAADVVSASITLQVLVDAPGNISCLWVKHSSLCQPHFD 119
Db 61 PEDLOCTPRQSGTVYEAATVEAESGSLTQVLATPGDLSCLWVKHSSLCQPHFD 120
QY 120 LQNRGVVSVILKMTQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQ 179
Db 121 LQNRGIVSMALLNVTEQAGEYLLHIOSEAAANYTLFTVNRDTQYVLRPRYFRKMENQ 180
QY 180 DALVCISESPEPIVEWVLCDSGECCKEESPAVVKKEKVLHELFGTDIRCCARNALGR 239
Db 181 DALLCISEGPEPTVEWVLCSSHRESCKEESPAVVKKEKVLHELFGTDIRCCARNALGR 240
QY 240 ECTRLFTIDLNTQPTLTPOLFLKVGEPWIRKAVHNVHGFLTWELNKALEEGNYFE 299
Db 241 ESTKLEFTIDLNQAPSTLTPOLFLKVGEPWIRKAVHNVHGFLTWELNKALEEGNYFE 300
QY 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVCKGFINATNSSE 359
Db 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTCSSSKHPSQSALVTIVCKGFINATNSSE 360
QY 360 EIDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHOFGEYIFH 419
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQEGLEDGYISKFCDHKNKGEYIFY 420
QY 420 AENDDAQFTKMTFLNIRRRKQVLAESAASQAFSDGYPLPSWTWKCDKSPNCTEIT 479
Db 421 AENDDAQFTKMTFLNIRKAPQVLANASQAFSDGYPLPSWTWKCDKSPNCTEIP 480
QY 480 EGVNRKANRKFVGQWVSSTLNMSEAIKGLVKKCCAYNSLGTSCETILNLSPPGPFPIQ 539
Db 481 EGVNKNKANRKFVGQWVSSTLNMSEAGLLVKKCCAYNSMGTSCTIFLNSPGPFPIQ 540
QY 540 DNIIFYATIGVLLFTVITLLCHYKKOFRYESQLOMVOVGTSSDNEYFYVDFREY 599
Db 541 DNIIFYATIGVLLFTVITLLCHYKKOFRYESQLOMVOVGTSSDNEYFYVDFREY 600
QY 600 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGSKTGVSIQAVKMLKADSSERELM 659
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGSKTGVSIQAVKMLKADSSERELM 660
QY 660 SELKMTQLGSHENIVNLGACTLSGPIYLFYCCYGDLNLYLRKREKFRHTWEIFK 719
Db 661 SELKMTQLGSHENIVNLGACTLSGPIYLFYCCYGDLNLYLRKREKFRHTWEIFK 720
QY 720 EHNFSYPTFQSHPNSSMPGSRVQIHPSDQISGLHGNFSHSEDEIYENQKRL--EEE 777
Db 721 EHNFSYPTFQSHPNSSMPGSRVQIHPSDQISGLHGNFSHSEDEIYENQKRL--EEE 780
QY 778 EDNLVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDPLGLARDIM 837
Db 781 EDNLVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDPLGLARDIL 840
QY 838 SDSNVYVRGNARLPVKWMAPELSLFEIGIYIKSDVWSYGILLMEIFSLGNVPYGPVVDAN 897
Db 841 SDSNVYVRGNARLPVKWMAPELSLFEIGIYIKSDVWSYGILLMEIFSLGNVPYGPVVDAN 900
QY 898 FYKLLQNGFKMDQPPYATEEIIYIMOSWAFDSRKRPSFNLTSFGLCCOLADAEAMQYN 957
Db 901 FYKLLQNGFKMDQPPYATEEIIYIMOSWAFDSRKRPSFNLTSFGLCCOLADAEAMQYN 960
QY 958 VDRGYSECPHTYQNRPRPSREMDLGLLSPQAQVE 991
Db 961 MGVNVPHEPSIYQNRPRPSREMDLGLLSPQAQVE 994

RESULT 15

PCT-US95-03718-2
; Sequence 2, Application PC/TUS9503718
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 879PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US95-03718-2

Query Match 86.0%; Score 4533.5; DB 5; Length 1000;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVPLLVVFSAMIFGTITNODLPVVKCVLINHKNNDSSVGSXSSYPVMS 59
Db 1 MRALAQRSDRRLLLVVLSVMIETVNTODLPVVKCVLINSHENNGSSAGKPSRYMVRGS 60
QY 60 PEDLGCALRQSSGTVYEAADVVSASITLQVLVDAPGNISCLWVKHSSLCQPHFD 119
Db 61 PEDLOCTPRQSGTVYEAATVEAESGSLTQVLATPGDLSCLWVKHSSLCQPHFD 120
QY 120 LQNRGVVSVILKMTQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRRPYFRKMENQ 179
Db 121 LQNRGIVSMALLNVTEQAGEYLLHIOSEAAANYTLFTVNRDTQYVLRPRYFRKMENQ 180
QY 180 DALVCISESPEPIVEWVLCDSGECCKEESPAVVKKEKVLHELFGTDIRCCARNALGR 239
Db 181 DALLCISEGPEPTVEWVLCSSHRESCKEESPAVVKKEKVLHELFGTDIRCCARNALGR 240
QY 240 ECTRLFTIDLNTQPTLTPOLFLKVGEPWIRKAVHNVHGFLTWELNKALEEGNYFE 299
Db 241 ESTKLEFTIDLNQAPSTLTPOLFLKVGEPWIRKAVHNVHGFLTWELNKALEEGNYFE 300
QY 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVCKGFINATNSSE 359
Db 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTCSSSKHPSQSALVTIVCKGFINATNSSE 360
QY 360 EIDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHOFGEYIFH 419
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQAFPCQEGLEDGYISKFCDHKNKGEYIFY 420
QY 420 AENDDAQFTKMTFLNIRRRKQVLAESAASQAFSDGYPLPSWTWKCDKSPNCTEIT 479

Db

421

|||||AENDDAQTKMFTLNIRKKQVLANASASQASCSDDGYPPLPSWTWKKCSDRSPNCTEIP

480

Qy

480

EGVWNRKANRVFGQWSSSTLNMSIAIKGFLVKCCAYNSLGTSCETILLNSPGPFPIQ

539

Db

481

EGVWNRKANRVFGQWSSSTLNMSIAIKGFLVKCCAYNSLGTSCETILLNSPGPFPIQ

540

Qy

540

DNISFYATIGVCLLFIIVLTLICHKKKQFYESQLOMQVQVGTGSSDNEYFYVDFREY

599

Db

541

DNISFYATIGVCLLFIIVLTLICHKKKQFYESQLOMQVQVGTGSSDNEYFYVDFREY

600

Qy

600

DLKWEFFRENLEFGKVLGSGAFKVMNATAYGISTGTYSIOVAVKMLKADSSREALM

659

Db

601

DLKWEFFRENLEFGKVLGSGAFKVMNATAYGISTGTYSIOVAVKMLKADSSREALM

660

Qy

660

SELKMTQLGSHENIVNLGACTLSGPIYLIIEYCCYGDLLNYLRKREKFRHTWTEIFK

719

Db

661

SELKMTQLGSHENIVNLGACTLSGPIYLIIEYCCYGDLLNYLRKREKFRHTWTEIFK

720

Qy

720

EHNFSPYTFQSHPNSSMPGSRVQIHPDSDQISGLHGNSFHSEDEIEYENOKRL--EEE

777

Db

721

EHNFSPYTFQSHPNSSMPGSRVQIHPDSDQISGLHGNSFHSEDEIEYENOKRLAEEEE

780

Qy

778

EDLNVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIM

837

Db

781

EDLNVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIL

840

Qy

838

SDSNVYVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN

897

Db

841

SDSNVYVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN

900

Qy

898

FYKLIQNGFKMDOPFYATEEIIYIMQSWAFDSRKRPSFNLTSFLGCQADAEAEAMQON

957

Db

901

FYKLIQNGFKMDOPFYATEEIIYIMQSWAFDSRKRPSFNLTSFLGCQADAEAEAMQON

960

Qy

958

VDGRVSECPHTYQNRPRPSREMDLGLLSPQAOVE

991

Db

961

MGGNVPEHPSIYQNRRLPSREAGSEPPSPQAVK

994

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:34:40 ; Search time 24.5123 Seconds

(without alignments)
4017.505 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5274	100.0	993	10	US-09-919-408-4
2	5274	100.0	993	10	US-09-872-136-4
3	4429.5	84.0	992	10	US-09-919-408-2
4	4429.5	84.0	992	10	US-09-872-136-2
5	1266	24.0	975	9	US-10-192-867-2
6	1251	23.7	976	9	US-10-099-895-32
7	1251	23.7	976	9	US-10-192-867-4
8	1216.5	23.1	972	10	US-09-944-807-10
9	1166.5	22.1	1088	9	US-09-961-403-4
10	1166	22.1	1089	9	US-09-955-363-36
11	1166	22.1	1089	10	US-09-769-987-2
12	1166	22.1	1089	10	US-09-919-497-90
13	1166	22.1	1089	10	US-09-866-510-2
14	1163	22.1	1089	10	US-09-866-510-10
15	1162	22.0	1089	10	US-09-866-510-4
16	1161	22.0	1089	10	US-09-866-510-8
17	1160	22.0	1089	10	US-09-866-510-6
18	1086.5	20.6	1090	10	US-09-866-510-14
19	1086.5	20.6	1106	9	US-09-955-363-2

ALIGNMENTS

RESULT 1

US-09-919-408-4

; Sequence 4, Application US/09919408

; Patent No. US20020072077A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTOPTENT HEMATOPOIETIC STEM CELL

; RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ImClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/919,408

; FILING DATE: 31-Jul-2001

; CLASSIFICATION: <Unknown>

; APPLICATION DATA:

; APPLICATION NUMBER: 07/977,451

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US 07/906,397

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: US PCT/US92/05401

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: TW 81102961

; FILING DATE: 15-APR-1992

; APPLICATION NUMBER: US PCT/US92/02750

; FILING DATE: 02-APR-1992

; APPLICATION NUMBER: US 07/813,593

; FILING DATE: 24-DEC-1991

; APPLICATION NUMBER: US 07/793,065

; FILING DATE: 15-NOV-1991

; APPLICATION NUMBER: US 07/728,913

; FILING DATE: 28-JUN-1991

20 1086.5 20.6 1106 10 US-09-866-510-22
21 1082.5 20.5 1106 10 US-09-866-510-16
22 1081.5 20.5 1106 10 US-09-866-510-20
23 1080.5 20.5 1106 10 US-09-866-510-18
24 990 18.8 1338 9 US-10-059-585-44
25 957 18.1 386 9 US-09-939-833-6
26 957 18.1 386 9 US-09-939-754-6
27 957 18.1 386 9 US-09-939-832-2
28 951.5 18.0 1356 9 US-10-022-939-2
29 951.5 18.0 1356 9 US-10-100-405A-2
30 950.5 18.0 1356 9 US-09-969-037-7
31 934 17.7 1367 10 US-09-766-678-2
32 923 17.5 1367 10 US-09-919-408-6
33 923 17.5 1367 10 US-09-872-136-6
34 907.5 17.2 1363 9 US-09-375-248-19
35 904.5 17.2 1298 10 US-09-982-610-33
36 904.5 17.2 1363 9 US-09-375-248-2
37 902.5 17.1 1368 9 US-10-105-901-34
38 902 17.1 1362 9 US-10-105-901-33
39 880.5 16.7 1363 9 US-10-105-901-32
40 859.5 16.3 367 9 US-09-939-833-9
41 859.5 16.3 367 10 US-09-939-754-9
42 859.5 16.3 367 10 US-09-939-832-9
43 846 16.0 367 9 US-09-939-833-12
44 846 16.0 367 10 US-09-939-754-12
45 846 16.0 367 10 US-09-939-832-12

Sequence 22, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 18, Appl
Sequence 44, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 19, Appl
Sequence 9, Appl
Sequence 33, Appl
Sequence 2, Appl
Sequence 34, Appl
Sequence 33, Appl
Sequence 32, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 12, Appl
Sequence 12, Appl

APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Query Match 100.0%; Score 5274; DB 10; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLNHNKNDSSVGKSSSPYMWSESP 60
Db 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLNHNKNDSSVGKSSSPYMWSESP 60

Qy 61 EDLGCALRPQSGTGYEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
Db 61 EDLGCALRPQSGTGYEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120

Qy 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMNED 180
Db 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMNED 180

Qy 181 ALVCISEVPEPIVWVLCDSOGESCKEPPAVVKEEKLVELFCTDIRCCARNELGRE 240
Db 181 ALVCISEVPEPIVWVLCDSOGESCKEPPAVVKEEKLVELFCTDIRCCARNELGRE 240

Qy 241 CTRFLTIDLNPOTTLPLQLFKVGEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300
Db 241 CTRFLTIDLNPOTTLPLQLFKVGEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300

Qy 301 STYSNRTMIRILFAVSSVANRDGYTCSSSKHPSQSALVTIVGKGINATNSSEYD 360
Db 301 STYSNRTMIRILFAVSSVANRDGYTCSSSKHPSQSALVTIVGKGINATNSSEYD 360

Qy 361 IDQEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYIFHA 420
Db 361 IDQEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYIFHA 420

Qy 421 ENDDAQFTKMTLINIRKPOVLAEASASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480
Db 421 ENDDAQFTKMTLINIRKPOVLAEASASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480

Qy 481 GYVNRKANRKFQWVSSSTLANSEAIKFLYKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GYVNRKANRKFQWVSSSTLANSEAIKFLYKCCAYNSLGTSCETILLNSPGPFPIQD 540

Qy 541 NTSFYATIGVCLLFIIVVTLILCHVKYKQFQRYESQWQVQVTCSSDNEYVDFREYED 600
Db 541 NTSFYATIGVCLLFIIVVTLILCHVKYKQFQRYESQWQVQVTCSSDNEYVDFREYED 600

Qy 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATYGISKTGVSQIYAVKMLKEKADSSERIALMS 660
Db 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATYGISKTGVSQIYAVKMLKEKADSSERIALMS 660

Qy 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIIFECYCGDLLNLYRSRKEKFRHTWTEIFKE 720
Db 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIIFECYCGDLLNLYRSRKEKFRHTWTEIFKE 720

Qy 721 HNFSTFPTQSHFNSSMPGSRVQIHPDSQISGLHGNFSFHSDEIEYENOKRLEEDL 780
Db 721 HNFSTFPTQSHFNSSMPGSRVQIHPDSQISGLHGNFSFHSDEIEYENOKRLEEDL 780

Qy 781 NVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

Qy 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNPNYPGIPVDANFYK 900

Qy 901 LIQNGFKMDQPFYATEEYIIMQSCWAFDSRKRPFNPNTLSGLCOLADAEEMAYQNVDG 960
Db 901 LIQNGFKMDQPFYATEEYIIMQSCWAFDSRKRPFNPNTLSGLCOLADAEEMAYQNVDG 960

Qy 961 RVSECPHTYQNRPFPSREMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPFPSREMDLGLLSPQAQVEDS 993

RESULT 2
US-09-872-136-4
; Sequence 4, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-872-136-4

Query Match 100.0%; Score 5274; DB 10; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVICKVLINHKNDSSVGKSSSYPMVWSESP 60
DB 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVICKVLINHKNDSSVGKSSSYPMVWSESP 60

QY 61 EDLGCALRPOSSGTYEAAAEVDVSAITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTYEAAAEVDVSAITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120

QY 121 QNRGVVSWILKMTETQAGYLLFIQSEATNYTILFTVSIIRNTLLYTLRPPYRKWENQD 180
DB 121 QNRGVVSWILKMTETQAGYLLFIQSEATNYTILFTVSIIRNTLLYTLRPPYRKWENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240

QY 241 CTRLEFTIDLNOTPOTTLPOLFLKVGEPWLIRCAVHVNHGFGLTWLENKALEEGNYFEM 300
DB 241 CTRLEFTIDLNOTPOTTLPOLFLKVGEPWLIRCAVHVNHGFGLTWLENKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIVGKGFNATNSSEDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIVGKGFNATNSSEDEYE 360

QY 361 IDQYEFCEFSVREKAYPOIRCTWTFGRKGFPCBQKGLDNGYSISKFCNHRKHQGEYIFHA 420
DB 361 IDQYEFCEFSVREKAYPOIRCTWTFGRKGFPCBQKGLDNGYSISKFCNHRKHQGEYIFHA 420

QY 421 ENDDAOFTRKFTLNIRRKPOVLAEASASQSCFSDGYPLPSWTWKCSKSPNCTEITE 480
DB 421 ENDDAOFTRKFTLNIRRKPOVLAEASASQSCFSDGYPLPSWTWKCSKSPNCTEITE 480

QY 481 GVMNRKANRKFVQWVSSSTLANSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GVMNRKANRKFVQWVSSSTLANSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIIVLTLILCHYKKQFRYESQLQWQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVLTLILCHYKKQFRYESQLQWQVGTSSDNEYFYVDFREYED 600

QY 601 LKWEPPRENLERFGKVLGSAFGKVMNATAYIGISKTGVSQVAVKMLKEKADSSEREA 660
DB 601 LKWEPPRENLERFGKVLGSAFGKVMNATAYIGISKTGVSQVAVKMLKEKADSSEREA 660

QY 661 ELKMMTQLGSHENIVNLGACTLSGPIYLIFECYCCGDLNLYLRSRREKPHRTWTIFKE 720
DB 661 ELKMMTQLGSHENIVNLGACTLSGPIYLIFECYCCGDLNLYLRSRREKPHRTWTIFKE 720

QY 721 HNFSEYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDETEYENOKRLEEEEDL 780
DB 721 HNFSEYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDETEYENOKRLEEEEDL 780

QY 781 NVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840
DB 781 NVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840

QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900

QY 901 LIQNGFRMDQFPYATEEYIIIMQSCWAFDSRKRPSFPNLTSLFLGCOLADAEAMYNQYVDG 960
DB 901 LIQNGFRMDQFPYATEEYIIIMQSCWAFDSRKRPSFPNLTSLFLGCOLADAEAMYNQYVDG 960

QY 961 RVSECPHTYQNRPRFSREMDLGLLSPOAQVEDS 993
DB 961 RVSECPHTYQNRPRFSREMDLGLLSPOAQVEDS 993

RESULT 3

US-09-919-408-2
Sequence 2, Application US/09919408
Patent No. US20020072072A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: InClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-919-408-2

Query Match 84.0%; Score 4429.5; DB 10; Length 992;
Best Local Similarity 84.1%; Pred. No. 1.7e-288;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

Qy 1 MPALA-RDAGTVPVLLVVSAMIFGTITNODLPVVKVCLVNHKNDSSVSKSSYPWMS 59
Db 1 MRALAQRSDRRLLLVLSVMTFVTTNODLPVVKVCLVSHENNSSAGKPSYRMVRS 60
Qy 60 PEDLGALRQSSGTVYEAADVEVDVSAITLQVLVDAPGNISCLWFKHSLNQCOPHD 119
Db 61 PEDLOCTPRQSGTVYEAATVEVAESGITLQVLATPGDLSCVWFKHSLNQCOPHD 120
Qy 120 LQNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPRYFRKMNQ 179
Db 121 LQNRGIVSMAILNVETQAGEYLLFIQSERANYTLFTVNVRTQLYLRRYFRKMNQ 180
Qy 180 DALVCISVPEPIVWVLCDSQSGESKPEEPAVVKKEVHLHFGTDIRCARNELGR 239
Db 181 DALLCISGVPEPIVWVLCDSHRESKEEGPAVVKKEVHLHFGTDIRCARNALGR 240
Qy 240 ECTRLFTIDLNOTPOTLPQLFKVGEPLWIRCKAVHNVHGFGLTWELENKALEEGNYE 299
Db 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIVHNVHGFGLTWELEDKALEEGSYE 300
Qy 300 MSTYSTNRTMIRILFAVSSVARNDTGYTCSKSPHPSQSALVTIVGKFINATNSSEY 359
Db 301 MSTYSTNRTMIRILFAVSSVGRNDGYTCSKSPHPSQSALVTILEKGFINATNSSEY 360
Qy 360 EDQYEEECFSVRFKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNHKHOPGEYIF 419
Db 361 EDQYEEECFSVRFKAYPIRCTWTFSRKSPCEOKGLDNGYSISKFCNHKHOPGEYIF 420
Qy 420 AENDDAQFTKMTLNRKPOVLAASASQSCFSDGYPLPWSWTWKCKSDKSPNCTEIT 479
Db 421 AENDDAQFTKMTLNRKPOVLANASASQSCSDGYPLPWSWTWKCKSDKSPNCTEIP 480
Qy 480 EGVWNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539
Db 481 EGVWNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 540
Qy 540 DNISYATIGVCLLFTVLTLLCHYKQFYESQLOVMQVGTGSSDNEFYVDFREY 599
Db 541 DNISYATIGVCLLFTVLTLLCHYKQFYESQLOVMQVGTGSSDNEFYVDFREY 600
Qy 600 DLKWFPPRENLEFGKVLGSGAFKVMNATAYIGKTVSIOQAVKMLKADSSREALM 659
Db 601 DLKWFPPRENLEFGKVLGSGAFKVMNATAYIGKTVSIOQAVKMLKADSSREALM 660
Qy 660 SELKMTQLGSHENIVNLGACTLSGPYILIFCYCYGDLNLYLRKREKFRHTWTEIFK 719
Db 661 SELKMTQLGSHENIVNLGACTLSGPYILIFCYCYGDLNLYLRKREKFRHTWTEIFK 720
Qy 720 EHNFSYPTFQSHPNSSMPCSRVQIHPDSDQJSLGNSHSEDETEYENOKRL--EEE 777
Db 721 EHNFSYPTFQSHPNSSMPCSRVQIHPDSDQJSLGNSHSEDETEYENOKRL--EEE 780
Qy 778 EDNLVLTFFDLICFAVOAKGMEFLFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIM 837
Db 781 EDNLVLTFFDLICFAVOAKGMEFLFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIL 840
Qy 838 SDSNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPYGPVVDAN 897
Db 841 SDSNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPYGPVVDAN 900
Qy 898 FYKLQNGKMDOPFTATEIYIIMOSCAFDKSRKPSFNLTSFLGCCLADAEAEAYQN 957
Db 901 FYKLQSGFKMEQFPFATYIYVMSOSCAFDKSRKPSFNLTSFLGCCLADAEAEAC--- 957
Qy 958 VDRGVSECPHYQNRPFREMDLGLLSPOAQVE 991
Db 958 ----IRTSIHLPKQAAPQQRG-GLRAQSPORQV 986

RESULT 4
US-09-872-136-2
; Sequence 2, Application US/09872136
; Patent No. US20020119545A1

GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-872-136-2

Query Match 84.0%; Score 4429.5; DB 10; Length 992;
Best Local Similarity 84.1%; Pred. No. 1.7e-288;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;
Qy 1 MPALA-RDAGTVPVLLVVSAMIFGTITNODLPVVKVCLVNHKNDSSVSKSSYPWMS 59
Db 1 MRALAQRSDRRLLLVLSVMTFVTTNODLPVVKVCLVSHENNSSAGKPSYRMVRS 60
Qy 60 PEDLGALRQSSGTVYEAADVEVDVSAITLQVLVDAPGNISCLWFKHSLNQCOPHD 119
Db 61 PEDLOCTPRQSGTVYEAATVEVAESGITLQVLATPGDLSCVWFKHSLNQCOPHD 120
Qy 120 LQNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPRYFRKMNQ 179
Db 121 LQNRGIVSMAILNVETQAGEYLLFIQSERANYTLFTVNVRTQLYLRRYFRKMNQ 180
Qy 180 DALVCISVPEPIVWVLCDSQSGESKPEEPAVVKKEVHLHFGTDIRCARNELGR 239
Db 181 DALLCISGVPEPIVWVLCDSHRESKEEGPAVVKKEVHLHFGTDIRCARNALGR 240
Qy 240 ECTRLFTIDLNOTPOTLPQLFKVGEPLWIRCKAVHNVHGFGLTWELENKALEEGNYE 299
Db 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIVHNVHGFGLTWELEDKALEEGSYE 300
Qy 300 MSTYSTNRTMIRILFAVSSVARNDTGYTCSKSPHPSQSALVTIVGKFINATNSSEY 359
Db 301 MSTYSTNRTMIRILFAVSSVGRNDGYTCSKSPHPSQSALVTILEKGFINATNSSEY 360
Qy 360 EDQYEEECFSVRFKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNHKHOPGEYIF 419
Db 361 EDQYEEECFSVRFKAYPIRCTWTFSRKSPCEOKGLDNGYSISKFCNHKHOPGEYIF 420
Qy 420 AENDDAQFTKMTLNRKPOVLAASASQSCFSDGYPLPWSWTWKCKSDKSPNCTEIT 479
Db 421 AENDDAQFTKMTLNRKPOVLANASASQSCSDGYPLPWSWTWKCKSDKSPNCTEIP 480
Qy 480 EGVWNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539
Db 481 EGVWNRKANRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 540
Qy 540 DNISYATIGVCLLFTVLTLLCHYKQFYESQLOVMQVGTGSSDNEFYVDFREY 599
Db 541 DNISYATIGVCLLFTVLTLLCHYKQFYESQLOVMQVGTGSSDNEFYVDFREY 600
Qy 600 DLKWFPPRENLEFGKVLGSGAFKVMNATAYIGKTVSIOQAVKMLKADSSREALM 659
Db 601 DLKWFPPRENLEFGKVLGSGAFKVMNATAYIGKTVSIOQAVKMLKADSSREALM 660
Qy 660 SELKMTQLGSHENIVNLGACTLSGPYILIFCYCYGDLNLYLRKREKFRHTWTEIFK 719
Db 661 SELKMTQLGSHENIVNLGACTLSGPYILIFCYCYGDLNLYLRKREKFRHTWTEIFK 720
Qy 720 EHNFSYPTFQSHPNSSMPCSRVQIHPDSDQJSLGNSHSEDETEYENOKRL--EEE 777
Db 721 EHNFSYPTFQSHPNSSMPCSRVQIHPDSDQJSLGNSHSEDETEYENOKRL--EEE 780
Qy 778 EDNLVLTFFDLICFAVOAKGMEFLFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIM 837
Db 781 EDNLVLTFFDLICFAVOAKGMEFLFKSCVHRDLAARNVLTGHKVKVVICDFGLARDIL 840
Qy 838 SDSNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPYGPVVDAN 897
Db 841 SDSNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPYGPVVDAN 900
Qy 898 FYKLQNGKMDOPFTATEIYIIMOSCAFDKSRKPSFNLTSFLGCCLADAEAEAYQN 957
Db 901 FYKLQSGFKMEQFPFATYIYVMSOSCAFDKSRKPSFNLTSFLGCCLADAEAEAC--- 957
Qy 958 VDRGVSECPHYQNRPFREMDLGLLSPOAQVE 991
Db 958 ----IRTSIHLPKQAAPQQRG-GLRAQSPORQV 986

```
121 LONRGIVSMALNVETQAGEYLLHIQSERANYTVLFTVNRDTQLVLRPFRKMEQ 180
QY 180 DALVCISEPPEIWEVLCDSGESCCKESPAPVKEEVHLHFLGTDIRCCARNELGR 239
Db 181 DALLCISEGPEPTVEWLCSSHRECKEKPAPVKEEVHLHFLGTDIRCCARNALGR 240
QY 240 ECTRLFTIDLNPQTLPOLFLKVGEPILWIRCKAVHNVHGFGLTWELNKALEEGNYFE 299
Db 241 ECTKLTIDLNOAPQSTLPOLFLKVGEPILWIRCKAVHNVHGFGLTWELNKALEEGNYFE 300
QY 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIYVKGFINATNSSEY 359
Db 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 360 EIDYEEFCFSVRFKAYPQIRCTWTFSRSPCEQKGLDNGYSISKFCNKHQGPYIFH 419
Db 361 EIDYEBKFCFSVRFKAYPQIRCTWTFSRSPCEQKGLDNGYSISKFCNKHQGPYIFY 420
QY 420 AENDDAQFTKMFNLNTRRKPOVLAERASASQSCFSGYPLPSPWTWKKCDKSPNCTEIT 479
Db 421 AENDDAQFTKMFNLNTRRKPOVLAERASASQSCFSGYPLPSPWTWKKCDKSPNCTEIP 480
QY 480 EGVNWKANRKYGVQWSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQ 539
Db 481 EGVNWKANRKYGVQWSSSTLNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
QY 540 DNISFYATTGCVLLFIVLWLLCHYKKQFRYESQLOMVOVQVGTSSDNEYFYVDFREY 599
Db 541 DNISFYATTGCVLLFIVLWLLCHYKKQFRYESQLOMVOVQVGTSSDNEYFYVDFREY 600
QY 600 DLKWEFPRENLEFGKVLGSCAFKGMNATAYGSKTGVSIOVAVKMLKADSSERALM 659
Db 601 DLKWEFPRENLEFGKVLGSCAFKGMNATAYGSKTGVSIOVAVKMLKADSSERALM 660
QY 660 SELKMTQOLGSHENIVNLLGACLSGPIYILFEYCCYDGLLNLYLSKREKFRHTWTEIFK 719
Db 661 SELKMTHLGHNDINVLGACLSGPIYILFEYCCYDGLLNLYLSKREKFRHTWTEIFK 720
QY 720 EHNFSYPTFQSHNPSMPSREVRQTHPSDQISGLHNSFHEDETEYENQKRL--EEE 777
Db 721 EHNFSYPTFQSHNPSMPSREVRQTHPSDQISGLHNSFHEDETEYENQKRLABEEE 780
QY 778 EDNLVLTFFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIM 837
Db 781 EDNLVLTFFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIL 840
QY 838 SDSNVYVRGNARLPVKWMAPESLFEGIYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 897
Db 841 SDSNVYVRGNARLPVKWMAPESLFEGIYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
QY 898 FYKLIONGFKMDOPFYATBEIYIIMOSCAFDSCRKPSFNLSFLGCCQLADAEAMQON 957
Db 901 FYKLIONGFKMDOPFYATBEIYIIMOSCAFDSCRKPSFNLSFLGCCQLADAEAEAC-- 957
QY 958 VDRGVSECPHTYQNRFPFREMDDLGLLSPQAQYE 991
Db 958 ----ITSHLPKQAAPQORG--GLRAQSPQRQVK 986
```

RESULT 5

```
US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-10-192-867-2

Query Match 24.0%; Score 1266; DB 9; Length 975;
Best Local Similarity 33.2%; Pred. No. 1.2e-76;
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

QY 35 CVLINHNRNDSSVCKSSSYPMVSESPDLGALRQSSGTVYEAAYVEVDVSASITLQVL 94
Db 12 CVLLVLLRGOTATQSPASGEPSP-----SIHPAQ-----ELIVEAGDTL--- 54
QY 95 VDAPGNLSCL-----VVEKHSSSLNCPHFDLQNRGVVSWILKMTQAGEYLLFIOSE 148
Db 55 -----SLTICIDPDFVRWTFK-----TYFN-----EMVENKKNEWIQ--EKAE 89
QY 149 ATNYTILFTVSIIRNTL-----LYTLRRPYFERKMNODALVCISESVPEPIVE 195
Db 90 ATR-TGYTCSNSGLTSSIIYFVRDPAKFLVLGLFLGK-EDSDALVRCPLTDPQ-VSN 146
QY 196 WVLCDGSGESCKESPAY-----VKKEKVHLEFGDIRCCARNELGRECTL--- 244
Db 147 YSLTECDGKSLPTDLTFVNPNGAGITIKNVRAYHRLC---VRCAAQ-----RDGTWLHSD 199
QY 245 -FTIDLNQT---PQTLPLQ--FLKYGEPLWIRCKAVHNVHGFGLTW-----E 286
Db 200 KFTLVKREAIKAIPIVSVSPETSHLLKKGDTFTVCTIKDVSTSVNSMWLKNMPOQHIAQ 259
QY 287 LENKALEEGNYFEMSTYTNTRMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVT--- 343
Db 260 VKHNSWRHGDG---NYERQETLF-----ISSARVDSGVFMCYANTTFSANVTTLK 309
QY 344 IVGKGFNAIN-SEDEYEDQYEEFCFSVRKAYPQ-IRCTWTFSRKSFQCEQKGLD--- 398
Db 310 VVEKGFNISPVKNTTFVTVDGENVDLVVEYAYPKPHQOQIYMNRT--SANKKDYVK 367
QY 399 -NGYSISKFCNKHK-----QPCGYIFHAENDDAQFTKMTLNRIRKPOVLA--EASAS 448
Db 368 SDNKSNIYVNLRLTRKLTGEGTYTFLVNSDASASVTENYVNTKPEILTYDRLING 427
QY 449 QACFSFGYPLPSPWTWKKCDKSPNCTEITE-GVWNKANKRVFGQWVSSSTLNSEAI 507
Db 428 MLQCVABEGPEPTIDWYFCIGABQRCITPVSPDVQVQNVSVSPGKLVQSSIDSSYFR 487
QY 508 KGFVLCAYNSLGTSCETILLNSPGPFPP---IQDN-----ISFYATTGCVLLFTIV 557
Db 488 HNGTVECKASNDVGKS--SAFFN---FAFKEQIQAHFLTPLTLIGFVVAAGAMGIIVNV 541
QY 558 LTLILCHYKKQFRYESOLOMV-OVTGSSDNEYFVDFREYEDLKWEFPRENLEFGKVL 616
Db 542 LT-----YKLOKPMYEQWKKVEEING---NNYIYIDTQPLPDHKKWEPFRNRLSFGKTL 594
QY 617 GSGAFGRVNMATAYGISKTGVSIQVAVKMLKEKADSSREALMSKMTOLGSHENIVN 676
Db 595 GAGAFGRVVEATAYGLIKSDAANTVAVKMLKPSAHLTEREALMSKLVLSLGNHMNI 654
QY 677 LLGACTLSGPIYILFEYCCYDGLLNLYLSKREKF-----HRTWTEIFKEHNFSPYTFQ 730
Db 655 LLGACTVGGPTLVITEYCCYDGLLNLRKRKDSIFTSKQEEQAEAAALYKLLHSTEPSCD 714
QY 731 SHPNSSM--PGSREVOIHPSDQISGLHNSFHEDEIEYENOKRLEEBEDNLVLTED 787
Db 715 S-SNEYMDMKPGYSYV-VPTKTKDRRSARIDSY-----IERDVTPIAMEDDEL-ALDD 766
QY 788 LLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVICDFGLARDIMSDSNVVRGN 847
Db 767 LLFSYQVAKAMAFASKNLCIHRDLAARNILTHGRITKICDFGLARDIRNDSNVVRGN 826
```

```

QY      848   ARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGPVDANFYKLIONGFK 907
         |||||:|||||::||| :||||| |||::||| :|||||::|||::|||::|||: |||: |||:
Db      827   ARLPVKWMAPEISFCVYYTFESDVWSYGIFLWFELSLGLSSPYPGMPVDSKCYKKKEGFR 886
         |||||:|||||::||| :||||| |||::||| :|||||::|||::|||::|||: |||: |||:
QY      908   MDOPFYATEEIIYMOSWAFSRKRSPFNLTSLFGCOLADAEAMYONV 958
         | :| :|||:|::||| :|||::| :|||::| :|||::| :|||::| :|||::| :|||::|
Db      887   MVSPEHAPEMYDMVTCTDADPLKRPTEKVQVVQLIEKIQISDSTRKIYSNL 937
         | :| :|||:|::||| :|||::| :|||::| :|||::| :|||::| :|||::| :|||::|

RESULT 6
US-10-099-895-32
; Sequence 32, Application US/10099895
; Patent No. US20020177166A1
; GENERAL INFORMATION:
; APPLICANT: BERNDT, Michael C
; APPLICANT: STONSKI, Frank C
; APPLICANT: LOPEZ, Angel F
; APPLICANT: GUTHRIDE, Mark A
; TITLE OF INVENTION: A BINDING MOTIF OF A RECEPTOR
; FILE REFERENCE: 3991/OK379
; CURRENT APPLICATION NUMBER: US/10/099,895
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/AU00/01118
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 976
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-099-895-32
```

	Query Match	23.7%	Score 1251;	DB 9;	Length 976;
	Best Local Similarity	32.7%;	Pred. No. 1.2e-75;		
	Matches 324;	Conservative 166;	Mismatches 352;	Indels 148;	Gaps 34;
Qy	47	VGKSSYPWSESPE	DLGALRPOSSGV	VEAAAEV	VDVSASITLQVLVDAPGNISCLWV 106
Db	20	VOTGSSQPSVFG	-EPSPSPH	PKSDLI	-----VRVDEIRLLCTDGFVK--WT 67
Qy	107	FKHSLNCQPHFD	LQNRGW	SMVILKMT	TOAGEYLLFIOSEATN---YTILFTVVSINT 163
Db	68	FE-----	ILDETENKQ	NEWIT-EKAEAT	NTGKYTCTNKHGLSNS 106
Qy	164	L-----	LYTLRRPFR	KMENODAL	VCISPEPIVENWVLCDSOGESCKE----- 209
Db	107	IYVFRDPAKFL	VDRLSYGK	-EDNDTIL	VRCLTPDPE-VTNYSLKGCQKPLKDLRIP 164
Qy	210	---SPA	VKKEKVLH	ELFGTD	IRCARNELGRECTRLFTIDL---NOTPOTLLPO--L 260
Db	165	DPKAGIMTK	VKRAYHRLC	---LHCS	VDQEGKSVLSEKFIKVRPAFRAVPVVSYSKASY 221
Qy	261	FLKVGEP	LWIRKAV	VHNGFGL	TWELEN---KALEEGNYFEMSYSYNRNWRILFAFV 317
Db	222	LLREGE	EFTVTCT	IKDVSSV	SYTWKRENSQTKLOEKYNWSHHGDFNVERQAT-----LTI 277
Qy	318	SSVARND	GYTTCSS	KHPSALVT	---IVCKGFNA-----TNSSEYDEIDQYE 365
Db	278	SSARVND	SGVFCY	ANNTFGS	ANVTTLTLEVDKDGFINFPMTNTVFVNDGENVDL---- 333
Qy	366	EFCFSVR	KAYPO-IR	CTWTF	SRKSF--PCEQKGLDNGYSIKFCNKHK-----QPGE 415
Db	334	----I	VEYAFK	PEHQWI	YNRRFTDKWEDYPKSENESNIRYVSELHLTKLGTGEGT 389
Qy	416	YIFHA	ENDDAQFT	KMFTL	LNIRKPOVLA--EASASQACFSFGYPLPSTWTKKSDKSPN 473
Db	390	YTFVLS	NSDVNA	IAAFNV	VYNTKPEILLTDRLVNGLQCVAAAGFPEPTIDWYFCPQTEOR 449
Qy	474	CTEETE	-GWN	RKANR	KVFGQWVSSTLNMSEAIKGLVKCAYNLSLGTSETILLNSP 532
Db	450	CSASVLP	VDQTLNS	SGPPG	FKLVQSSIDSFAFKHNGTVECKAYNDVGKT--SAYFN-- 505

```

RESULT 7
US-10-192-867-4
; Sequence 4, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS-002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-192-867-4

Query Match      23.7%; Score 1251; DB 9; Length 976;
Best Local Similarity 32.7%; Pred. No. 1.2e-75;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps

Qy 47 VGKSSSYPMVSESPEDLCALRPOSSGTVYEAAREVDVVSASITLQVLVDAPGNISCLW 106
Db 20 VQTGSSQSPVSPG-EPSPSPHPGKSDLI-----VRVGDEIRLLCTDPPGVK--WT 67
Qy 107 FKHSSLNCQPHFDLQNRGVSMVILKMTETQAGEYLLFIOSEATN---YTILETVSIRNT 163
Db 68 FE-----ILDENENKQNEWIT-EKAEATNTCKYCTCTNKHGLSNS 106
Qy 164 L-----LYTLRRPFRKMEMODALVCISESVPEPPIVWVLCDSQGSECKE----- 209
Db 107 IYVVRDPAKFLVDVRSIYK-EDNDTLVRCLPTDPE-VTNYSLKGCGRPLPKDLRFIP 164
Qy 210 ---SPAVYKKEEVLHFGTDIRCCARNELGRECTRLFTIDL-----NOTPQTTLPO--L 260

```

Query Match	23.7%;	Score 1251;	DB 9;	Length 976;
Best Local Similarity	32.7%;	Pred. NO. 1.2e-75;		
Matches 324:	Conservative 166;	Mismatches 352;	Indels 148;	Gaps 34;

Qy	47	VGRSSSYPMVSESPEDLGALRPQSSGTVYAAAEVDVSATITQLVLDAPGNISCLMW	106
Db	20	VQTGSSQPSVSPG-EPSPPSIIHGKSDLI-----VRVGDEIRLLCTDGPGEV- -WT	67
Qy	107	FKHSSLNCQPHFDLQNRGVVSMILKMTQAGEYLLFTIOSEATN---YTILFTVSVIRMT	163
Db	68	FE-----ILDETENKQEWIT-EKAEATNGKYTCTNKHGLSNS	106
Qy	164	L-----LYTLRRPYPRKMNODALVCISGSPVEPIVWVLCDSQSGSCKEE-----	209
Db	107	IYVVRDPAKFLVDRSLYCK-EDNTLVRCPTDPE-VTNYSLKCCGCKPLPKDLRFIP	164
Qy	210	---SPAVYKKEEVLHFGTDIRCCARNELGRECTRLFTIDL-----NOTPQTTLPO--L	260

Db 165 DPRAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFIKVRPAKPAKVPVSVSKASY 221
QY 261 FLAVGEPLAIRKAVHVNHCGLTWELN---KALEEGNYFEMSTYTNRTMIRILFAFV 317
Db 222 LUREGEEFTVTIKDVSSSVYTWKRENSOTLQEKYNSWHHGDGFYERQAT---LTI 277
QY 318 SSVARNDTGYTCCSSKHPQSALVT---IVRGKFINA-----TNSDEYEDQYE 365
Db 278 SSARVNDGVEMCYANNITFGSANVTITLEVVDKGFINFIPMINTVTVFVNDGENVDL---- 333
QY 366 EFCFSVRFAKPO-IRCTWTFPSKSF--PCEQKGLDNGYSISKFCNKH-----OPGE 415
Db 334 ---IVEYEAFFPKPEHQOMIYMNRTETDKWEDYPKSESNENIYVSELHLTRLKGTGGT 389
QY 416 YIFHAENDDAQFTKMTLNIARRKPOVLA---EASASQSCFSDGYPPLPSWTWKCKSDKSPN 473
Db 390 YTLVNSDVNAIAFNVYVNTKPEILTYDRLVNGMLQCVAAGFPEPIDIMYFCPGTEQR 449
QY 474 CTEEITE-GVMNRKANRKFQGVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSP 532
Db 450 CSASVLPVDVQTLNLSGPPGKLVVSSIDSSAFKHNGTVECKAYNDVGT--SAYFN-- 505
QY 533 GPPFPFIQDN-----ISFYATIGVCLLFIWLITLLCHIKYKKQFYRESOLQM 578
Db 506 --FAPGNNKEQIHPHTLFTPLLIGFVIVAGMCIIMILT---YKYLQKPMVEVQWKV 559
QY 579 V-QVTGSSDNEFYDFREYEDLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTGV 637
Db 560 VEEING--NNYVIIDPTQLPYDHKWEFFPRNLSFGKTLGAGAGKVVATEYGLIKSDA 616
QY 638 SIQVAVKMLKEKADSEREALMSLKMMTOLGSHENIVNLLGACTLSGPYLLIFEYCCYG 697
Db 617 AMTVAVKMLKPSAHLTERALMSLKVSLYLGNNHMINVLLGACTIGGPTLVITEYCCYG 676
QY 698 DLLNLYRSKREKFRHTWTE-----IFKEHNFSEYPTFQSHPNSSM---PGSREVOIHPD 748
Db 677 DLLNFLRRKRSDFICSKQEDHAEALYKLLHSEKSSCDSTNEYMDKPGSVV-VPTK 735
QY 749 SDQISGLHNSHSEDEIEYENOKRLEEBEDLNLVTFEDLLCFAYQVAKGMEFLEPKSCV 808
Db 736 ADKRRSVRIGSY-----IERDVTAPAMEDEL-ALDLEDLLSFYSYQAKGMFLASKNCI 789
QY 809 HRDLAARNVLTGHGVKVCICDGLARDIMSDSNVYVGNARLVPKMAPESLFEGIYTIK 868
Db 790 HRDLAARNLTLTHGRTIKCDGLARDIKNDSNVYVGNARLVPKMAPESIFNCVTFE 849
QY 869 SDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWAF 928
Db 850 SDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWAF 909
QY 929 DSKRPSFNLTSFLCQCLADAEAEANYQNV 958
Db 910 DPLKRPTFQIOVLIEKQISESTNHIYSNL 939

RESULT 8

US-09-944-807-10
; Sequence 10, Application US/09944807
; Patent No. US20020119494A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim Pharma KG
; TITLE OF INVENTION: Method for identifying substances which positively
; TITLE OF INVENTION: Influence inflammatory conditions of chronic
; TITLE OF INVENTION: Inflammatory airway diseases
; FILE REFERENCE: 082.00n
; CURRENT APPLICATION NUMBER: US/09/944,807
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: UK 0021484.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10

LENGTH: 972
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-944-807-10

Query Match 23.1%; Score 1216.5; DB 10; Length 972;
Best Local Similarity 32.8%; Pred. No. 2.5e-73;
Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

QY 75 VYEAAYEVDDV--SASITLQVLVDAPGNISCLWFKHSSLNCQPHFDLONRGVSWVILK 132
Db 22 VIEPSVPELVYKPGATVLRVCV---GNGSVEMDGPSS---PHWTLYSDGSSILSTLN 72
QY 133 MTE-----TCAGEYL-----LFIOSEATNYTILFTVTSIRNTLLYTLRRRYFERKM 176
Db 73 NATQNTGTVTRCTEPGDPGLGSAAILHLYVKDAPRWNVL-----AOEVVVF----- 118
QY 177 ENQDALV-CISESPERPIEIVWLCDQSGESCKEESPAVVKKEK-----VLRE--- 223
Db 119 EQDALLPCL--LTDVPLE-----AGVSLVRVGRPLMRHTNYSFSPWHGFTIHRAKF 169
QY 224 LFGTDIRCCARNELGRECTRLFTIDLNOTPOTTLP-----OLFLKVGEEPLWIRCK 273
Db 170 IQSDYQCSALMG-GR---KVMISIRLKVQKVIKPPALTLVPAELVIRGEAAQIVCS 225
QY 274 AVHNVHFGTLWELENKALEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDTGYTCCSS 333
Db 226 ASSVDVNFDFVLOHNNTKLA---IPQOSDFHNNRYQ-KVLTLNLDQVDFQHAGNYSVAS 281
QY 334 ---KHPQSQALVTIVCKGFINATNNSD---YEIDQVEEFCFSVRFAKAYPOIR-CTWTF 385
Db 282 NVQCKH-STSMFFRVVESAYLNL--SSEONLQIEVTVEGLNLKVVWEAYGLQGFNTY 338
QY 386 SRKSFPCQKGLDNGYSISKFNHQP-----GEXIF 418
Db 339 -----LGPSDHPPEKLANATTKTDTYRHTFTLSLPLAKPSEAGRYSF 381
QY 419 HAENDDAQFTKMTLNIARRKPOVLAESAASOAS-----CFSDGYPLPSTWTKCKSDKSPNC 474
Db 382 LARPPGWRALTFTLTRYPEVSVIWTFFINGSGTLLCAASGYPOPNTVTLQCSGTHDRC 441
QY 475 TEEITEGVN---RKANRKFVGVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLN 530
Db 442 DEAQVLQVDDPYEVLDSQEPFHVKTVQSLTLETLEHNTQYECRAHNSVSGSWAFIPI 501
QY 531 SPGPFPIQDNISFYATIGVC---LLFIVTLTLLCHIKYKKQFYRESOLQVQVVTGSSDN 587
Db 502 SAGAHTHPDDEFLTTPVVVACMSIMALLLLLLLLKYLKQKPKYQVRWKIIE--SYEGN 559
QY 588 EYFVVDREYEDLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOAVKMLK 647
Db 560 SYTFIDPTQLPYNEKWEFFPRNNLQFGKTLGAGAFGVVEATAFGLGEDAVLKVAVKMLK 619
QY 648 EKADSSEREALMSLKMMTOLGSHENIVNLLGACTLSGPYLLIFEYCCYGDLLNLYRSKR 707
Db 620 STAHADEKALMSLKMIMSHLQHENIVNLLGACTHGGPVLVITEYCCYGDLLNLYRSKR 679
QY 708 E-----KFHRTWTEIFKEHNFSEYPTFQSHPNSSMPCGREVOIHP 747
Db 680 EAMLGPSLSPQDPEGGVDYKNIHLEKKYVVRDGSF-----SQGVDIYV 724
QY 748 DSQISGLHNSHSEDEIEYENOKRLEEBEDLNLVTFEDLLCFAYQVAKGMEFLEPKSC 807
Db 725 EMRPVSTSSNDSF-SEQDLKEDGRPLE-----LRDLLHFSOYAOQMAFLASKNC 774
QY 808 VHRDLAARNVLTGHGVKVCICDGLARDIMSDSNVYVGNARLVPKMAPESLFEGIYTI 867
Db 775 IHRDLAARNVLTGHGVKVCICDGLARDIMSDSNVYVGNARLVPKMAPESIFDCVTV 834
QY 868 KSDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWA 927
Db 835 QSDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWA 894

REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-955-363-36

Query Match 22.1%; Score 1166; DB 9; Length 1089;
Best Local Similarity 30.3%; Pred. No. 6.8e-70;
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNQLPVI-----KCVLINHKNDSSVSKSS---SYPMVSES 59
DB 6 PAFVLGCLLTGLSLILCOLSLPNEKVVQLNSFSRLRCFGESEVSWQYPMSEEE 65

QY 60 PEDLGCALRQSSG---TVYEAAAVEVDVSATITLOVLVDAPGNISCLWFKHSSLCQP 116
DB 66 SSDVEIRNEENNSGLFVTIVLEVSS-----ASAHT-----GLYTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTQAGEYLLFTIOSEATNTILFTVSIRNTLLTYL 168
DB 111 --ELEGRIIYVDPDPAFVPLGHT-----DYLIVVE----- 141

QY 169 RRPYFRKMNODALVCISESVPEPIVEWVLCDSSG-----ESCKE 208
DB 142 -----DDDSAIIPCTTDPETPV--TLHNSGVVPASVDSRQGFNGFTVGPYIC-- 189

QY 209 ESPAVYKKEKVLHLEFGDIRCARNELGRECTRLFTIDLQTPQTLFOLFKVGEPL 268
DB 190 EATVKGKKFTQIPFNYYALK-----ATSELDMEALKTV-----YKSGETI 231

QY 269 WIRCKAVHNHGFGLTW-----ELENA-----LEGNYPFEMSTYSTNRTMIRILFAF-VSSV 320
DB 232 VVTC-AVFNEVVVDLOWTYGEVKGKGITILE-----IKVPSIKLVYTLTVP 280

QY 321 ARNDTYTCSSK-----HPSOALVTIVGKFINATNSSEDEY-IDOYEFCFSVRFK 374
DB 281 TVKDSGDYECARQATREVKEMKVVITSVHEKGFIEIKPTFSOLEAVNLHVKHFVVEVR 340

QY 375 AYPQIRCTWTFSEKSPCEQKGLDNGYSIS-----KFCNKHQ-PGEYIF 418
DB 341 AYPPIPRISLKNLTL-----IENLTETITDVEKIOEIRYSKLLIRAKEEDSGHYTI 394

QY 419 HAENDDAQFTKMTLNRKPOVL-----ABASASOASCFSDDGYPLPSWTKKCSDKS 471
DB 395 VAQNEADAVKSYTFELLTOVPSSILDLVDHGGSTGGQTVRCTAEGTLPDIEWMICKD-I 453

QY 472 PNCTEEITEGVNMRKANRKFVGOWSS-----STLNMSAIGFLVKCCAY 517
DB 454 KCKNNETS---WILANN-----VSNIIIEHSRORSTVEGRVTFPAKVEETAVRCLAK 504

QY 518 NSLGTSCETILLNSPGPPFIQNIISFYATIGVCLFIVVLTLLCHIKYKQRYESOLQ 577
DB 505 NLLGAENRELKVA-----PTLRSELTAVAALVLLVITVILSVLVVWIKQPKRYEIRWR 560

QY 578 MVQVTGSSNEXEYVDFREYDLKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGV 637
DB 561 VIESISPDGHEIYVDPMLPYDSRWEFFRGLVGLVGLGSGAFGKVEGTATGLSRSQP 620

QY 638 SIQVAYKMLKEKADSSEREALMSLKMWTOLGSHENIVNLLGACTLSGGPIYLIFEYCCY 697
DB 621 VMKVAVKMLKPTARSSEKQALMSLKMTHLGPLNIVNLLGACTSGPIYIITEYFCY 680

QY 698 DLLNLYRSKREK-----HRTWTEIFKEHNFSPYFTFQSHPN 734
DB 681 DLVNLHKNRDSFLSHRPEKPKKELDIFGLNPADESTRSVILSFENNMDYMDMKQADTT 740

RESULT 11
US-09-769-987-2
Sequence 2, Application US/09769987
Patent No. US20020055129A1
GENERAL INFORMATION:
APPLICANT: Matsui, Toshimitsu
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Recept
FILE REFERENCE: 14014. 0266U2
CURRENT APPLICATION NUMBER: US/09/769,987
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 08/460,656
PRIOR FILING DATE: 1995-06-02
PRIOR APPLICATION NUMBER: US 08/439,095
PRIOR FILING DATE: 1995-05-11
PRIOR APPLICATION NUMBER: US 07/915,884
PRIOR FILING DATE: 1992-07-20
PRIOR APPLICATION NUMBER: US 07/308,282
PRIOR FILING DATE: 1989-02-09
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1089
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1e -
OTHER INFORMATION: synthetic construct
US-09-769-987-2

Query Match 22.1%; Score 1166; DB 10; Length 1089;
Best Local Similarity 30.3%; Pred. No. 6.8e-70;
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNQLPVI-----KCVLINHKNDSSVSKSS---SYPMVSES 59
DB 6 PAFVLGCLLTGLSLILCOLSLPNEKVVQLNSFSRLRCFGESEVSWQYPMSEEE 65

QY 60 PEDLGCALRQSSG---TVYEAAAVEVDVSATITLOVLVDAPGNISCLWFKHSSLCQP 116
DB 66 SSDVEIRNEENNSGLFVTIVLEVSS-----ASAHT-----GLYTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTQAGEYLLFTIOSEATNTILFTVSIRNTLLTYL 168
DB 111 --ELEGRIIYVDPDPAFVPLGHT-----DYLIVVE----- 141

QY 169 RRPYFRKMNODALVCISESVPEPIVEWVLCDSSG-----ESCKE 208
DB 142 -----DDDSAIIPCTTDPETPV--TLHNSGVVPASVDSRQGFNGFTVGPYIC-- 189

```
QY 209 ESPAVYKKEKVLHFLGTDIRCCARNELGRECTRLFTIDLNQTPOTTLPOLFKVGEPL 268
Db 190 EATVKGKKFQTIIPFNVAALK-----ATSELDLEMEALKT-----YKSGETI 231
QY 269 WIRCKAVVHNGFGLTW-----ELENKA-----LEEGNYFEMSTYTNRTMIRILFAF-VSSV 320
Db 232 VVTC-AVFENNEVDLQWTPGEVKGKGITMLEE-----IKVPSIKLVTLTLPVEA 280
QY 321 ARNDTGYTCSSSK-----HPSQSALVTIVGKGFINATNSSEDEYE-IDQYEEFCFSVRFK 374
Db 281 TVKDSGDYECARQAATREVEMKVVITISVHEKGFIEIKPTFSQLEAVNLHEVKEHVVVEVR 340
QY 375 AYPQIRCTWTFESKSPCEQKGLDNGYSIS-----IENLTEITDVEKIOEIRYRSKLIRAKEEDSGHYTI 418
Db 341 AYPPPRISLWKNLTL-----PTLRSELTVAAAVLVLLVIISLVLVVWVKOPRYEIRWR 394
QY 419 HAENDDAQTKMFTLNIRKRPQVL-----AEASASQSCFSDGYPPLSWTWKKSCKS 471
Db 395 VAQNEADAVKSYTFELLTQVPSIILDDVDDHGGSTGGTQVTRCTAEGTPLPDIEWMICKD-I 453
QY 472 PNCTEEITGVVNRKANRKFVQWVSS-----STLNMSEAIGFLVKCCAY 517
Db 454 KCCNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGFPFFIODNISFYATIGVCLLFIIVVLTLLICHKKQRYESQLQ 577
Db 505 NLLGAENRELKVA-----PTLRSELTVAAAVLVLLVIISLVLVVWVKOPRYEIRWR 560
QY 578 MVQVTGSSDNEYFYDFREYEDLKWEPFRENLEFGKVLGSGAFKGMVNNATAYGISKTGV 637
Db 561 VIESISPDGHEIYVDPQPLPYDSRWEPFRDGLVRLGVLGSGAFKGMVNNATAYGISKTGV 620
QY 638 SIOVAVKMLKEKADSRREALMSELKMMTQLGSHENIVNLLGACTLSGPIYLIFECYCG 697
Db 621 VMKVAVKMLKPTARSSEKQALMSELKIMTHLGPLHINVLNLLGACTKSGPIYIIEYCYG 680
QY 698 DLNLYRSKRKEF-----HRTWTEIFKEHNFSTYFTFQSHPN 734
Db 681 DLVNYLHKNRDSFSLSHHPEKPKKELDFGLNPADESTRSYVILSFENNDDYMDMKQADTT 740
QY 735 SSMGSRVQIHPDSQOISGLHGN---SPHSEDEIYEYNQKRLKEEEDLNVLTFEDLLCFA 792
Db 741 QYVPMLEKKEVSKYSQISQSLYDRPASYKKKMLDSE-VKNLLSDDDNSEGLTLLDLSFT 799
QY 793 QYVAKGMEFLFKSCVHRDLAARNVLVTHCKVKVICKDFGLARQIMSDSNVYVGRNARLPV 852
Db 800 QYVARGMEFLASKNCVHRDLAARNVLLAOKKIVKICDFGLARQIMSDSNVYVGRNARLPV 859
QY 853 KWMAPESLPGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVDPANFYKLIQNGFKMDQPF 912
Db 860 KWMAPESIFDNLTYTISDWSYGILLWEIFSLGTPYPGMVDSTFYNKIKSGYRMAKPD 919
QY 913 YATEIYIMQSCAFDRKRPPNLTSPFGCOLADAEAMYNQV---DGRVSECPH--- 967
Db 920 HATSEVYEIMVKWCNSEPEKRPSPFYLHSEIVENLLPGQYKKSEKIHLDFLKSHPAVAR 979
QY 968 -----TYQNRPPRSREMDICL 983
Db 980 MRVDSONAYIGVTYKNEEDKLKDWEGGL 1007
```

RESULT 12

```
US-09-919-497-90
; Sequence 90, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Muttter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919, 497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
```

```
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-90

Query Match      22.1%   Score 1166;   DB 10;   Length 1089;
Best Local Similarity 30.3%;   Pred. No. 6.8e-70;
Matches 336;   Conservative 165;   Mismatches 365;   Indels 242;   Gaps 36;
```

```
QY 12 PLLVYFSAMIFG---TITNODLPVI-----KCVLIHKNNDSSVKGSS---SYPMVSES 59
Db 6 PAFVLGCLLTGLSLILCOLSLPSILPNEKEKVVQLNSSFSLRCFGESEVSWQPMSEEE 65
QY 60 PEDLGCALRPOSSG---TVYEAAAVEVDVSASITQLVLDVADPAGNISCLWFKHSLNCP 116
Db 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAAHT-----GLYTCYINHQTQEN--- 110
QY 117 HFDLQNRGV-----VSMVILKMTETQAGEYLLFIQSEATNTVILFTVSIIRNTLLYTL 168
Db 111 --ELEGRIIYVDPDVAFLPGMT-----DYLIVB----- 141
QY 169 RRPYFRKMNODALVCISESYPEPIVEMVLCDSOG-----ESCKE 208
Db 142 -----DDSAIIPCRTTDPETPV---TLHNSGVVVPASYDSRGGFNGFTVGVPIYC- 189
QY 209 ESPAVYKKEKVLHFLGTDIRCCARNELGRECTRLFTIDLNQTPOTTLPOLFKVGEPL 268
Db 190 EATVKGKKFQTIIPFNVAALK-----ATSELDLEMEALKT-----YKSGETI 231
QY 269 WIRCKAVVHNGFGLTW-----ELENKA---LEEGNYFEMSTYTNRTMIRILFAF-VSSV 320
Db 232 VVTC-AVFENNEVDLQWTPGEVKGKGITMLEE-----IKVPSIKLVTLTLPVEA 280
QY 321 ARNDTGYTCSSSK-----HPSQSALVTIVGKGFINATNSSEDEYE-IDQYEEFCFSVRFK 374
Db 281 TVKDSGDYECARQAATREVEMKVVITISVHEKGFIEIKPTFSQLEAVNLHEVKEHVVVEVR 340
QY 375 AYPQIRCTWTFESKSPCEQKGLDNGYSIS-----KFCNHRKHQ-PGEYIF 418
Db 341 AYPPPRISLWKNLTL-----IENLTEITDVEKIOEIRYRSKLIRAKEEDSGHYTI 394
QY 419 HAENDDAQTKMFTLNIRKRPQVL-----AEASASQSCFSDGYPPLSWTWKKSCKS 471
Db 395 VAQNEADAVKSYTFELLTQVPSIILDDVDDHGGSTGGTQVTRCTAEGTPLPDIEWMICKD-I 453
QY 472 PNCTEEITGVVNRKANRKFVQWVSS-----STLNMSEAIGFLVKCCAY 517
Db 454 KCCNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGFPFFIODNISFYATIGVCLLFIIVVLTLLICHKKQRYESQLQ 577
Db 505 NLLGAENRELKVA-----PTLRSELTVAAAVLVLLVIISLVLVVWVKOPRYEIRWR 560
QY 578 MVQVTGSSDNEYFYDFREYEDLKWEPFRENLEFGKVLGSGAFKGMVNNATAYGISKTGV 637
Db 561 VIESISPDGHEIYVDPQPLPYDSRWEPFRDGLVRLGVLGSGAFKGMVNNATAYGISKTGV 620
QY 638 SIOVAVKMLKEKADSRREALMSELKMMTQLGSHENIVNLLGACTLSGPIYLIFECYCG 697
Db 621 VMKVAVKMLKPTARSSEKQALMSELKIMTHLGPLHINVLNLLGACTKSGPIYIIEYCYG 680
QY 698 DLNLYRSKRKEF-----HRTWTEIFKEHNFSTYFTFQSHPN 734
Db 681 DLVNYLHKNRDSFSLSHHPEKPKKELDFGLNPADESTRSYVILSFENNDDYMDMKQADTT 740
QY 735 SSMGSRVQIHPDSQOISGLHGN---SPHSEDEIYEYNQKRLKEEEDLNVLTFEDLLCFA 792
Db 741 QYVPMLEKKEVSKYSQISQSLYDRPASYKKKMLDSE-VKNLLSDDDNSEGLTLLDLSFT 799
QY 793 QYVAKGMEFLFKSCVHRDLAARNVLVTHCKVKVICKDFGLARQIMSDSNVYVGRNARLPV 852
```

Db 800 YQVARGMEFLASKNCVHRDLAARNVLLAQAQKIVKICDFGLARDIMHDSNVYSGSTLPV 859
QY 853 KWMAPESLEGITIKSDVMSYGILLWEIFSLGVNPPGIPVDANFYKLLQNGFKMDQPF 912
Db 860 KWMAPESIFDNLTSLSDVMSYGILLWEIFSLGCTPYPGMMVDSTFYNNKISGYRMAKPD 919
QY 913 YATEEYIIIMQSWAFDSRRKPSFPNLTSLGCOLADAEAMYNV--DGRVSECPH--- 967
Db 920 HATSEVYEMVKNWSEPEKPSFYHLSEIVENLLPGQYKSKYEKIHLDPLKSDHPAVAR 979
QY 968 -----TYQNRPPFSREMDLGL 983
Db 980 MRVDSNAYIGVTYKNEEDKLKDWEGGL 1007

RESULT 13
US-09-866-510-2
; Sequence 2, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-866-510-2

Query Match 22.1%; Score 1166; DB 10; Length 1089;
Best Local Similarity 30.3%; Pred. No. 6.8e-70;
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVVFSAIFG---TITNQDLPIV-----KCVLINHKNDSSVGKSS---SYPMVSES 59
Db 6 PAFLVGLCLTGLSLILCOLSPLIPNENKVVQNLNFSFLRCFGESEVSWQYPMSEEE 65
QY 60 PEDLGCALRPQSSG---TYVEAAVEVDYSASITLQVLVDAPGNISCLMWFKHSSLNCQP 116
Db 66 SSDVEIRNEENSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110
QY 117 HFDLQNRGV-----VSMVILKWTETOAGEYLLFIQSEATNWTILFTVSIINTLLYTL 168
Db 111 --ELEGHRIYIYVDPDVAFLPGMT-----DYLIVIVE----- 141
QY 169 RRPYFRKNQDALVCISEVPEPIVEMVLCDSOG-----BSCKE 208
Db 142 -----DDSAIPCRITDDETPV--TLNSEGVVVPASYDSRQGFNGFTVGPIC-- 189
QY 209 ESPAVVVKKEKVLHELFGDIRCCARNELGRECTRLFTIDLNOTPTTLPLQLKVGCEPL 268
Db 190 EATVRGKKFQTIPIFNVAULK-----ATSELDLEMEALKTV-----YKSGETI 231
QY 269 WTRCAVHVNHGFLTW-----ELENA-----LEGNYFEMSTYNTNRWIRILFAF-VSSV 320
Db 232 VVTC-AVFENNEVVDLQWTPGVEKRGKITMLEE-----IKVPSIKLVTLTVPEA 280
QY 321 ARNDGYTCSSSK-----HPQSALVTIVGKFINATNSSEDIY-IDQVEECFVSREK 374
Db 281 TVKSGDVECAARQATREYKEMKKTIVSHEKGFIEIKPTFSQLEAVNLNHEVRFHVEVR 340
QY 375 AYPQIRCTWTFSRKSPFCQKGLDNGYSIS-----KFCNHHQV-PGEYIF 418

Db 341 AYPPPRISWLKNNLTL-----IENLTETTTVEKIQEIRYRSKLKLIKRAKEEDSGHVTI 394
QY 419 HAENDDAQFTKFTLNIRRKQDVL-----AASASOASCFSDCYPLPSTWTKCSDKS 471
Db 395 VAQNEADAVKSTFELLTQVPSSILDLVDHGHGSGTGGQTVRCATAGTPLPDIEMTICKD-I 453
QY 472 PNCFTTEITEGVWNRKANKRVGQWVSS-----STLNMSEAIGFLVKCCAY 517
Db 454 KCCNNETS---WTLANN-----VSNITIHSDRSTVEGRVTFFAKVEETIARVCLAK 504
QY 518 NSLGTSCETILLNSPGPPFPFIQDNISFYATIGVCLLFIIVLTLTLLCHYKXKOFRESIQ 577
Db 505 NLLGAENRELKLV-----PTLRSLETVAALVLLVIVILVILVIVIKOKPREIRWR 560
QY 578 MVQVTGSSDNEYFYDFREYEDLKNFPRENLEFGKVLGSGAGKVMNATAYGTSKTV 637
Db 561 VIESISPDGHEIYIYDPMQLPFDYSRWEFFPRDGLVLRVGLGSGAFGVVEGTAYGLSRQP 620
QY 638 SIOVAVKMLKEKADSSERREALMSKMTOLGSHENIVNLLGACTLSGPIVLIPEYCCYG 697
Db 621 VMKVAVRMLKPTARSEKQALMSKIMTHLGPHLNIVNLLGACTSGPIIITEYCYFG 680
QY 698 DLLNLRSKREKF-----HRTWTEIFKEHNFVSFYPTFQSHPN 734
Db 681 DLVNYLHKNRDSFLSHHPKPKKELDIFGLNPADESTRSYVILSFENNGDYMDMKQADTT 740
QY 735 SSMPGSEVOIHPDSQISGLHGN--SFHSEDEIEYENQKRLEEEEDNLVTLFEDLLGFA 792
Db 741 QYVPMLEKREKVSYSKYSDIORSYDRPASYSKKSMLDSE-VKNLLSDDNSEGLTLLDLSFT 799
QY 793 QYVAKGMEFEFKSCVHRDLAARNVLYTHGVKVICDFGLARDIMSDSNVYVGRNARLPV 852
Db 800 QYVARGMEFLASKNCVHRDLAARNVLLAQAQKIVKICDFGLARDIMHDSNVYSGSTFLPV 859
QY 853 KWMAPESLEGITIKSDVMSYGILLWEIFSLGVNPPGIPVDANFYKLLQNGFKMDQPF 912
Db 860 KWMAPESIFDNLTSLSDVMSYGILLWEIFSLGCTPYPGMMVDSTFYNNKISGYRMAKPD 919
QY 913 YATEEYIIIMQSWAFDSRRKPSFPNLTSLGCOLADAEAMYNV--DGRVSECPH--- 967
Db 920 HATSEVYEMVKNWSEPEKPSFYHLSEIVENLLPGQYKSKYEKIHLDPLKSDHPAVAR 979
QY 968 -----TYQNRPPFSREMDLGL 983
Db 980 MRVDSNAYIGVTYKNEEDKLKDWEGGL 1007

RESULT 14
US-09-866-510-10
; Sequence 10, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1089
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-866-510-10

Query Match 22.1%; Score 1163; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 1.1e-69;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVVSAMIFG---TITNODLPV1-----KCVLINHKNDSSVKGSS---SYPMVSES 59
Db 6 PAFLVLGCLTGLSLILCOLSLPILPNEKENKVVQNSFSLRFCGESEVSWQYPMSEEB 65
QY 60 PEDLGCALRPOSG---TVYEAADVVDVSASITLQVLVDAPGNISCLWFKHSSLNCOQ 116
Db 66 SDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110
QY 117 HFDLQNGV-----VSMVILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLTYL 168
Db 111 --ELEGRHIYIYVDPDVAFAVPLGTM-----DYLIVIVE----- 141
QY 169 RRPYFRKMNODALVCISSEVPEPIVENVLCDOSG-----ESCKE 208
Db 142 -----DODSAILPCRTTDPETPV--TLNSEGVPVPAVSYSRQGFNGFTVGPYIC-- 189
QY 209 ESPAVVVKKEKVLHELFGTDIRCCARNELGRECTRLFTIDLNQTPQTTLPLQLFKVGBPL 268
Db 190 EATVKGKKFQIPENVYALK---ATSELDLEMEALKTV-----YKSGETI 231
QY 269 WIRCKAVHNVHGFGLTW-----ELENK-----LEEGNYFEMSTYSTNRTMIRILFAP-VSSV 320
Db 232 VVTC-AVFNNVVDLQWTPGVEKVGKITMLEE-----IKVPSIKLVYTLTVPEA 280
QY 321 ARNDTGYTCCSSK-----HPSQSALVTIVGKGFINATNSSDYE-IDOYEERCFVSVPK 374
Db 281 TVKDSGDYECARQAOTREVKEMKVTISVHEKGFIEIKPTFSQLEAVNLHEVHFVVEVR 340
QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----IENLTETDVEKIQIRYRSKLLIRAKEEDSGHYTI 394
Db 419 HAENDDAQFTKMTLNIIRKPOVL-----AASASQACSFSDGVPPLSWTWKCKSDKS 471
Db 395 VAQNEADVKSFTFELLTQVPSILLDLDHGGSTGGTGTCTAEGTPLPDIEWMICKD-I 453
QY 472 PNCTEEITEGVWNRKANRKFVQGVVSS-----STLNNSEAIKFLVKCCAY 517
Db 454 KCKNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGPPFIQDNISFYATIGVCLLFIWVLTLLICHKKYKQFVRESQ 577
Db 505 NLLGAENRELKIVA---PILRSELTVAAAVLVILVILVILVIVWVKQPEYRWR 560
QY 578 MVQVGTSSDNEYFYDFREYEDLKNWEPRENLEFGVLGSGAPGYKYNATAYGISKTGV 637
Db 561 VIESISPDGHEIYIYVDPMLPYDSRWEFFRDLGLVGRVLGSGAPGVVEGTAYGLRSQP 620
QY 638 SIOVAVKMLKADSSEREAALMSLKMATOLGSHENIVNLLGACTLSGPYILIFECYCG 697
Db 621 VMKAVAVKMLKPTARSSKQALMSLKMATOLGSHENIVNLLGACTLSGPYILIFECYCG 680
QY 698 DLNLYLRSKREKF-----HRTWTEIFKEHNFSPYTFQSHPN 734
Db 681 DLNLYLHKNDLSLHHPKPKKELDIFGLNPADESTRSYVILSFENNQDYMMDKQADTT 740
QY 735 SSMPSGSEVOIHPSDQISGLHGN--SFHSEDEIEYENQKREBEEDNLVLTFFDLCLFA 792
Db 741 QYVPMLEKREKSVKSYDIQSLYDRPASYKKKMLDSE-VKNLLSDDNSEGLTLLDLST 799
QY 793 QYVAKGMEFLFKSCVDRDLAARNVITHGKVVKICDFGLARDIMSDSNVVRGNARLPV 852
Db 800 QYVARGMEFLASKNCHVRDLAARNVLAQKIVKICDFGLARDIMHDSNYSVSGSTFLPM 859
QY 853 KWAPESLFEIGIYTKSDVMSYGLLWEIFSLGNVPYGPVDPDANFYKLIQKMDOPF 912
Db 860 KWAPESIFONLYTLSDVMSYGLLWEIFSLGTPPGMVMWDTFYNNKIKSGYRMAKPD 919
QY 913 YATIEIYIMQSCWAFDRSRPSPNLTSLFGCOLADAEAEAMQNV--DGRVSECPH--- 967
Db 920 HATSEVTEIWMKCNSEPERKPSFYHLSEIIVENLLPGYKKYSKEIKIHLDFLKSHPAVAR 979

QY 968 -----TYONRRPFSREMDLGL 983
Db 980 MRVDSNAYIGVYTKNEEDKLKDWEGGL 1007
RESULT 15
US-09-866-510-4
; Sequence 4, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-4

Query Match 22.0%; Score 1162; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 1.3e-69;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVVSAMIFG---TITNODLPV1-----KCVLINHKNDSSVKGSS---SYPMVSES 59
Db 6 PAFLVLGCLTGLSLILCOLSLPILPNEKENKVVQNSFSLRFCGESEVSWQYPMSEEB 65
QY 60 PEDLGCALRPOSG---TVYEAADVVDVSASITLQVLVDAPGNISCLWFKHSSLNCOQ 116
Db 66 SDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110
QY 117 HFDLQNGV-----VSMVILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLTYL 168
Db 111 --ELEGRHIYIYVDPDVAFAVPLGTM-----DYLIVIVE----- 141
QY 169 RRPYFRKMNODALVCISSEVPEPIVENVLCDOSG-----ESCKE 208
Db 142 -----DODSAILPCRTTDPETPV--TLNSEGVPVPAVSYSRQGFNGFTVGPYIC-- 189
QY 209 ESPAVVVKKEKVLHELFGTDIRCCARNELGRECTRLFTIDLNQTPQTTLPLQLFKVGBPL 268
Db 190 EATVKGKKFQIPENVYALK---ATSELDLEMEALKTV-----YKSGETI 231
QY 269 WIRCKAVHNVHGFGLTW-----ELENK-----LEEGNYFEMSTYSTNRTMIRILFAP-VSSV 320
Db 232 VVTC-AVFNNVVDLQWTPGVEKVGKITMLEE-----IKVPSIKLVYTLTVPEA 280
QY 321 ARNDTGYTCCSSK-----HPSQSALVTIVGKGFINATNSSDYE-IDOYEERCFVSVPK 374
Db 281 TVKDSGDYECARQAOTREVKEMKVTISVHEKGFIEIKPTFSQLEAVNLHEVHFVVEVR 340
QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----IENLTETDVEKIQIRYRSKLLIRAKEEDSGHYTI 394
Db 419 HAENDDAQFTKMTLNIIRKPOVL-----AASASQACSFSDGVPPLSWTWKCKSDKS 471
Db 395 VAQNEADVKSFTFELLTQVPSILLDLDHGGSTGGTGTCTAEGTPLPDIEWMICKD-I 453
QY 472 PNCTEEITEGVWNRKANRKFVQGVVSS-----STLNNSEAIKFLVKCCAY 517
Db 454 KCKNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGPPFIQDNISFYATIGVCLLFIWVLTLLICHKKYKQFVRESQ 577

Db 505 NLGAENRELKVA---PTLRSELTVAAGVLLVLIIVISLIVVIMQKQPRYEIRWR 560
QY 578 MVQVTCSSNEIYFVDFREYEDLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGV 637
Db 561 VIESTISPDGHEIYVDPMLPYDSRWKFFPRDGLVGLGVLGSGAFGKYVECTAYGLSRSP 620
QY 638 SIQVAVKMLKEKADSSEREAALMSLKMOTGLSHENIVNLLGACTLSGPIYLIFEYCCYG 697
Db 621 VMKVAVKMLKPTARSEKQALMSLKMOTGLSHENIVNLLGACTLSGPIYLIFEYCCYG 680
QY 698 DLNLYLRKREKF-----HRTWTEIFKEHNSFYPTFQSHPN 734
Db 681 DLVNYLHKNRDSEFLSHHPEKPKKELDIFGLNPADESTRSVILSFENNNGDYMDMKQADTT 740
QY 735 SSMPSGREVOIHPDSQIISGLHGN--SFHSEDEIEYENQKRLEEBEDLNVLTFTEDLLCFA 792
Db 741 QYVPMLEKREKVSYSIQRSYDRPASYYKKKMLDSE-VKNLLSDDNSEGLTLDDLST 799
QY 793 YOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVVRGNARLPV 852
Db 800 YQVARGMEFLASKNCVHRDLAARNVLLAQKIVKICDFGLARDIMHDSNYSKGSFLPV 859
QY 853 KWMAPESEFEGYTIKSDVMSYIGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPF 912
Db 860 KWMAPESEFENLYTTLLSDVMSYIGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPF 919
QY 913 YATEEYIIMQSCWAFDSRKRPSFPMNLSFGLCOLADAEAMYQNV--DGRVSECPH--- 967
Db 920 HATSEYIEMVKWCNSEPEKRPSEFYLSEIVENLLPGQYKKSYEKIHLDPKSDHFAVAR 979
QY 968 -----TYQNRPPFSREMDLGL 983
Db 980 MRVDSDNAYIGVTKNEEDKLDWEGGL 1007

Search completed: May 27, 2003, 14:45:10
Job time : 29.5123 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:27:39 ; Search time 26.5134 seconds

(without alignments)
3600.506 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDLGLLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5238	99.3	993	2 A36873	protein-tyrosine k
2	4533.5	86.0	1000	2 S18827	Flt3 protein - mou
3	4429.5	84.0	992	2 A39931	protein-tyrosine k
4	1286	24.4	977	2 I45877	protein-tyrosine k
5	1266	24.0	975	1 TVMSKT	protein-tyrosine k
6	1263.5	24.0	978	1 A49814	protein-tyrosine k
7	1251	23.7	976	1 TVHUKT	protein-tyrosine k
8	1226	23.2	954	2 I51703	c-kit-related kina
9	1225.5	23.2	980	1 TVCTMD	macrophage colony-
10	1218	23.1	941	1 TVHUMD	protein-tyrosine k
11	1216.5	23.1	972	1 TVHUMD	macrophage colony-
12	1215	23.0	960	1 JN0677	protein-tyrosine k
13	1213	23.0	975	2 T30816	macrophage colony-
14	1194	22.6	978	2 S16385	macrophage colony-
15	1183.5	22.4	976	1 TVMSMD	macrophage colony-
16	1176	22.3	1088	1 PFTTGA	platelet-derived g
17	1166	22.1	1089	1 PFHUGA	platelet-derived g
18	1146	21.7	1089	1 S33727	platelet-derived g
19	1128	21.4	1087	2 S15522	platelet-derived g
20	1105.5	21.0	1098	1 PMSRB	platelet-derived g
21	1086.5	20.6	1106	1 PFHUGB	platelet-derived g
22	1080	20.5	1048	2 T30815	platelet-derived g
23	992	18.8	790	1 FQMVHZ	gag-kit polyprotei
24	990	18.8	1338	2 S09982	protein-tyrosine k
25	959.5	18.2	1333	2 S178875	receptor tyrosine
26	956	18.1	1336	2 I60598	Flt-1 tyrosine kin
27	944.5	17.9	1356	2 J61402	protein-tyrosine k
28	939.5	17.8	1330	2 S49010	embryonic receptor
29	935	17.7	1348	2 S51656	vascular endotheli

30	932	17.7	1379	2 JC4954	vascular endotheli
31	933	17.5	1367	2 A41228	protein-tyrosine k
32	907.5	17.2	1363	2 I58375	protein-tyrosine k
33	904.5	17.2	1298	2 A48999	protein-tyrosine k
34	874.5	16.6	823	2 B35963	protein-tyrosine k
35	851	16.1	160	2 A39061	protein-tyrosine k
36	849	16.1	821	1 TVHUF2	fibroblast growth
37	842	16.0	813	1 A49123	fibroblast growth
38	841	15.9	824	2 S24108	protein-tyrosine k
39	836.5	15.9	822	2 A45081	fibroblast growth
40	832.5	15.8	824	2 S36439	fibroblast growth
41	828.5	15.7	769	2 S16236	fibroblast growth
42	828.5	15.7	822	2 A41794	keratinocyte growt
43	827	15.7	797	2 S38579	fibroblast growth
44	825	15.6	821	1 TVMSBK	fibroblast growth
45	823.5	15.6	822	2 B54846	fibroblast growth

ALIGNMENTS

RESULT 1

A36873

protein-tyrosine kinase (EC 2.7.1.112), STK-1 precursor - human

N:Alternate names: stem cell tyrosine kinase 1

C:Species: Homo sapiens (man)

C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 04-Feb-2000

C:Accession: A36873

R:Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; B

Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994

A:Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+)

A:Reference number: A36873; MUID:94119906; PMID:7507245

A:Accession: A36873

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-993 <SMA>

A:Cross-references: GB:U02687

A:Note: In the authors translation, an additional residue Ala is shown after 420-Ala

C:Genetics:

A:Map position: 13q12

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-sp

F:608-950/Domain: protein kinase homology <KIN>

F:616-624/Region: protein kinase ATP-binding motif

Query Match 99.3%; Score 5238; DB 2; Length 993;

Best Local Similarity 99.7%; Pred. No. 2.3e-247;

Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNDSSVSKSSSPVMSVSESP 60

DB 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNDSSVSKSSSPVMSVSESP 60

QY 61 EDLGCALRPOSSGTVVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120

DB 61 EDLGCALRPOSSGTVVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120

QY 121 QNRGVVSWILKMTQTQAGEYLLFIQSEATNTYTLFVSIRNTLLYTLRRPYFKMKENQD 180

DB 121 QNRGVVSWILKMTQTQAGEYLLFIQSEATNTYTLFVSIRNTLLYTLRRPYFKMKENQD 180

QY 181 ALVCISSEVPPEVWVLCDSQGESCKEESPAVVKKEKVLHFLGFTDIRCCARNELGRE 240

DB 181 ALVCISSEVPPEVWVLCDSQGESCKEESPAVVKKEKVLHFLGFTDIRCCARNELGRE 240

QY 241 CTRLFITDINQTPQTTLPLQLFLKVGPELWIRCAVHVHNGFGLTWELNKALEBGNFEM 300

DB 241 CTRLFITDINQTPQTTLPLQLFLKVGPELWIRCAVHVHNGFGLTWELNKALEBGNFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNQDGYTTCSSSKHPQSQALVTITVGKFINATNSSDYE 360

DB 301 STYSTNRTMIRILFAFVSSVARNQDGYTTCSSSKHPQSQALVTITVGKFINATNSSDYE 360

A:Reference number: A39931; MUID:91292518; PMID:164848

A:Accession: A39931

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-992 <MAT>

C:Cross-references: GB:M64689; NID:q193327; PIDN:AAA37634.1; PID:q193328

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki

F:609-953/Domain: protein kinase homology <KIN>

F:617-625/Region: protein kinase ATP-binding motif

Query Match 84.0%; Score 4429.5; DB 2; Length 992;

Best Local Similarity 84.1%; Pred. No. 4.6e-208;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MPALA-RDAGTVPLAVFSAMIFGHTITQDLPVIRKVLINHNDSVSGKSSYPMVSS 59

DB 1 MRALAQRDRRLLLVLSVILETVTQDLPVIRKVLISHENNGSSACKPSSYMRVRS 60

QY 60 PEDLGALRPOSSGTVYAAAVEVDVSASITLOVLVDAPGNSICLWFKHSLNCPHFED 119

DB 61 PEDLOCTPRROSEGVYEAATVEAESGSIITQVOLATPGDLSCLWFKHSLGCGPHFD 120

QY 120 LQNRGVSMVLKMTQAGEYLLFIOSEATNYTILFTVSIIRNTLLYTURRPYFRKMQ 179

DB 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTFLTVNVVDLTQLVLRPYFRKMQ 180

QY 180 DALVCISFSEVPPIVWVLCDSQGSKEESPAAVVKKEKVLHELFGTDIRCARNELGR 239

DB 181 DALLCISEGVPETVWVLCSSHRESKEEGPAVVRKKEKVLHELFGTDIRCARNALGR 240

QY 240 ECTRFLTDLNQTPOTLPOLKLYGVEPLWIRKAVVNHGGLTWELNKALEEGNYFE 299

DB 241 ECTKLTFTDLNQAPOSTLPOLKLYGVEPLWIRKAIHNVHGFGLTWELDKALEGSYFE 300

QY 300 MSTYSTNTMIRILFAVSSVARNDTGYTCSSSKHPSOSALVTIVGKGFINATNSSEY 359

DB 301 MSTYSTNTMIRILFAVSSVARNDTGYTCSSSKHPSOSALVTILEKGFINATNSSEY 360

QY 360 EIDYEEFCFSVREKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHQPGGYIEH 419

DB 361 EIDPYEKFCFSVREKAYPRICTWTFWFSQASFPCEORGLDGYISKFCNKHQPGGYIEY 420

QY 420 AENDDAQTKMFTLNIIRKPOVLAAASASQASCFSDGYPLPSWTKKCDKSPNCTEIT 479

DB 421 AENDDAQTKMFTLNIIRKPOVLAAASASQASCFSDGYPLPSWTKKCDKSPNCTEIT 480

QY 480 EGVNKRANRVFGOMVSSSTLNMSEATKGLV KCCAYNSLGTSCETILLNSPGPPFIQ 539

DB 481 EGVNKRANRVFGOMVSSSTLNMSEAGKLLV KCCAYNSMTGTSCTIFLNSPGPPFIQ 540

QY 540 DNISFYATIGVCLLPFIVVLTLLCHIKKQFRYESOLOMVOVQVTSDDNEFYVDFREY 599

DB 541 DNISFYATIGVCLLPFIVVLTLLCHIKKQFRYESOLOMVOVQVTSDDNEFYVDFREY 600

QY 600 DLKWEFPRENLEFGVLGSGAFKVMNATAYGISTGVSIQVAVKMLKADSSEREALM 659

DB 601 DLKWEFPRENLEFGVLGSGAFKVMNATAYGISTGVSIQVAVKMLKADSSEREALM 660

QY 660 SELKMTOLGSHENIVNLGACTLSGPVILFEYCCYGDLNLYLRSKREKFRHTWTEIFK 719

DB 661 SELKMTOLGSHENIVNLGACTLSGPVILFEYCCYGDLNLYLRSKREKFRHTWTEIFK 720

QY 720 EHNFSYFTQSHPNSSMPGSEVQIHPDSDOI SGLHGNFSHSEDEIEYENOKRL--EEE 777

DB 721 EHNFSYFTQSHPNSSMPGSEVQIHPDSDOI SGLHGNFSHSEDEIEYENOKRLAEEB 780

QY 778 EDNLVLTEDLLCFAYQVAKMEFEFFKSCVHRDLAARNVLTGHKVKVLCDFGLARDIM 837

DB 781 EDNLVLTEDLLCFAYQVAKMEFEFFKSCVHRDLAARNVLTGHKVKVLCDFGLARDIL 840

QY 838 SDSNVYVRGNARLPVKWAPESLFEGIIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897

DB 839 SDSNVYVRGNARLPVKWAPESLFEGIIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 898

DB 841 SDSNVYVRGNARLPVKWAPESLFEGIIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900

QY 898 FYKLIONGKMDOPFYATEEIIYIMOSWAFDSKRKPSFNLTSFLGCLADAEAEAMQN 957

DB 901 FYKLIONGKMDOPFYATEEIIYIMOSWAFDSKRKPSFNLTSFLGCLADAEAEAC--- 957

QY 958 VDRGVSECHTIVNRRPFSREMDLGLLSPOAOYE 991

DB 958 ----IFTSIHLPKQAAPOQRG-GLRAQSPORQVK 986

RESULT 4

I45877

Protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs

C:Species: Bos primigenius (aurochs)

C>Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 20-Apr-2000

C:Accession: I45877

R:Kubota, T.; Hikono, H.; Sasaki, E.; Sakurai, M.

Gene 141, 305-306, 1994

A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.

A:Reference number: I45877; MUID:94215924; PMID:7512939

A:Accession: I45877

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-977 <KUB>

C:Cross-references: GB:D16680; NID:9516659; PIDN:BAA04084.1; PID:g516660

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

F:329-395/Domain: immunoglobulin homology <IMM>

F:588-932/Domain: protein kinase homology <KIN>

Query Match 24.4%; Score 1286; DB 2; Length 977;

Best Local Similarity 33.2%; Pred. No. 2.7e-55;

Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;

QY 47 VKGSSYPVMSPEPDLGC-ALRPQSSGTVYAAAVEVDVSASITLOVLVDAPGNISCLW 105

DB 20 VQTGSSQPSV--SPGELSUPSHPAKS-----ELIVSGDEIRLLCTDPGFVK-W 66

QY 106 VFKHSSLNCQPHDLQNRGVSMVLKMTQAGEYLLFIOSEAT---NYTILFTVSIRN 162

DB 67 TFE-----ILQOLSEKTNPEWIT-EKAEATNTGNTCTNKGLSS 105

QY 163 TL-----LYTLRRPYFRKMENODALVCISESPEPIVWVLCDSQGSKEESPAP 213

DB 106 SIYFVVRDPEKLELDLPLYGKEEN-DTLVRCPLTDPE-VTNYSLTGCCKPLPKDLTFV 163

QY 214 -----VKKEKVLHELFGTDIRC--CARNELGRE-CTRLETFID----LNQTPQTLP 258

DB 164 ADPKAGITIRNVKREYHRL-----CLHCSANRGKSMLSKKTFLKVRAAIKAVPVVSVS 217

QY 259 Q--LFLKVGEPWIRKAVVNHGGLTWELNKALEEGNVFEMSTYSTNRTMIRILFAF 316

DB 218 KTSYLLREGEFAVTCILIKDVSSVDSMWIKENSOQTAKQTKNNSHQGDFSYLRQBELT 277

QY 317 VSSVARNDTGYTCSSKHPQSALVT---IVGKGFINA-----TNSSDEYDIQY 364

DB 278 ISSARVNDSCVFCYANNTFGSANVTTLLEVVDKGFINIFPMNTTVFVNDGENVDL--- 334

QY 365 EEEFCFSVRKAYPO-IRCTWTFSRKS-----FPCQKGLDNGYSISKFCNKHK----- 411

DB 335 -----VVEEAYPKPVHRQWYINMRTSTDKWDYDPKSE-----NESNIRVYNELHLTRLK 384

QY 412 --OPGEYIFHAENDDAQTKMFTLNIRRKPOVLAAESA--SQASCFSDGYPLPSWTWKK 467

DB 385 GTEGGTYTHVNSDVNSSTVFNVYVNTKPEILTHDLVNGMLQCVAAAGPEPTIDWYFC 444

QY 468 SKSPNCTEII-TEGVNKRANRVKFGQWVSSSTLNMSEAIKGLFKKCCAYNSLGTSET 526

DB 445 PGTEQRCSPVGPVDVQIQNSVSPGKLVVYSTIDSTEFKHNGTVECRAYNDVGKSSAS 504

QY 527 ILLNSGPPPP-----FTQDNISFYATIGVCLLPFIVLTLCHIKKQFRYESOLOM 578

DB 528 ILLNSGPPPP-----FTQDNISFYATIGVCLLPFIVLTLCHIKKQFRYESOLOM 579

Db 505 FNFAFGKSGKEQIHAHTLTFTLLIGFVIAAGLMCIFVMILT-----YKYLQKPMYEVQWKV 560
Qy 579 V-QVTGSSDNEFFYDFREYEDLKWEPRENLEPGKVLGSCAFCKVMNATAYGISKTCV 637
Db 561 VEEING-----NNVYIDPTQPLFDHKKWEPNRLSEKTLGAGAFKGVVEATAYGLIKSDA 617
Qy 638 SIQVAVKMKKEKADSSEREAALMSLKMNTOLGSHENIYNLLGACTLSGPIYILIFYCCYG 697
Db 618 AMTAVVAKMKPSAHLTEREALMSLKLVSILGNHNNIYNLLGACTIGGPTLVITEYCCYG 677
Qy 698 DLLNVLRSKRKEFHRTWTE-----IFKEHNPFFPTTQSHPNSSM---PGSREVOIHPD 748
Db 678 DLLNPLRRKRDSPFCQKQEDHAEEVALYNLLHSKSSCNDSTNEYMDKMPGVSYV-VPTK 736
Qy 749 SDQISLGHNSFHSDELEYENOKLEEEEDNLVLTFFDLFCAYOVAKGMFLEPKSCV 808
Db 737 ADKRSARISG-----IERDVTPTAMEDDEL-ALDLELLSFQYVAKGMFLASKNCI 790
Qy 809 HRDLAARNLVTHGKVVKICDFGLARDMSDSYVVRGNARLPVKWMAPESEFEGIYTIK 868
Db 791 HRDLAARNILLTHGRITKICDFGLARDIKDSYVVKGNARLPVKWMAPESEIFNCYITEE 850
Qy 869 SDVWSYGILLWEIFSLGNVPYGPVDPANFYKLIQNGFKMDQPFYATEBIYIIMOSCWAF 928
Db 851 SDVWSYGIFLWELFSLGSSPYGPMVPVDSKFKWIKEGFRMLSPHAPADMIDMKTCDWA 910
Qy 929 DSRKRPSPNLTSLFCGLADAEAEAMYNQVGRVSEC-PH 967
Db 911 DPLKRPFTKQIVOLIEKQISESTNNIYSN-----LANCSPH 946
RESULT 5
TVMSKT
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - mouse
N;Alternate names: tyrosine kinase receptor c-kit
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 31-Mar-2000
C;Accession: S00474; B44876; I49596
R;Qiu, F.; Ray, P.; Brown, K.; Barker, P.E.; Jhanwar, S.; Ruddle, F.H.; Besmer, P.
EMBO J. 7, 1003-1011, 1988
A;Title: Primary structure of c-kit: relationship with the CSF-1/PDGF receptor kinase fa
A;Reference number: S00474; MUID:88296403; PMID:2456920
A;Accession: S00474
A;Molecule type: mRNA
A;Residues: 1-975 <QID>
A;Cross-references: GB:Y00864; NID:g50423; PIDN:CAA68772.1; PID:g50424
R;Rossi, P.; Marziani, G.; Albanesi, C.; Charlesworth, A.; Geremia, R.; Sorrentino, V.
Dev. Biol. 152, 203-207, 1992
A;Title: A novel c-kit transcript, potentially encoding a truncated receptor, originates
A;Reference number: A44876; MUID:92331813; PMID:1378413
A;Accession: B44876
A;Molecule type: DNA
A;Residues: 771-814 <ROS>
A;Note: sequence extracted from NCBI backbone (NCBIN:108837; NCBI:P108840)
R;Yasuda, H.; Galli, S.J.; Geissler, E.N.
Biochem. Biophys. Res. Commun. 191, 893-901, 1993
A;Title: Cloning and functional analysis of the mouse c-kit promoter.
A;Reference number: I49596; MUID:93221533; PMID:7682073
A;Accession: I49596
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-22 <RES>
A;Cross-references: GB:L11358; NID:g293325; PIDN:AAA37420.1; PID:g293326
C;Genetics:
A;Gene: kit; c-kit
A;Map position: 5
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
fic protein kinase
F.1-22/Domain: signal sequence #status predicted <SIG>
F.23-975/Product: protein-tyrosine kinase kit #status predicted <KTC>
F.23-519/Domain: extracellular #status predicted <EXT>
F.51-100/Domain: immunoglobulin homology <IMM>

F.130-189/Domain: immunoglobulin homology <IMM2>
F.227-295/Domain: immunoglobulin homology <IMM3>
F.331-397/Domain: immunoglobulin homology <IMM4>
F.426-496/Domain: immunoglobulin homology <IMM5>
F.520-542/Domain: transmembrane #status predicted <IMM>
F.543-975/Domain: intracellular #status predicted <INT>
F.586-929/Domain: protein kinase ATP-binding motif
F.594-602/Region: protein kinase ATP-binding motif
F.58-98.137-187.234-293.431-494/Disulfide bonds: #status predicted
F.146.296.303.323.355.370.466.489/Binding site: carbohydrate (Asn) (covalent) #status
F.622.639.790/Active site: Lys, Glu, Asp #status predicted
F.755.808/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 24.0%; Score 1266; DB 1; Length 975;

Best Local Similarity 33.2%; Pred. No. 2.5e-54;

Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

Qy 35 CVLINHKNDSVVGKSSSYPMVSESPEDLGCALRPOSSGTVYEAAYVVDVSASITLOVL 94

Db 12 CVLLVLLRGOTATSPSPGPPSP-----SIHPAQSS-----ELIVEAGDTL--- 54

Qy 95 VDPAGNISCL-----WVFKHSLNCPHFQDLQNRGVSMVLKMTQTQAGEYLLFIQSE 148

Db 55 -----SLTCIDPFDVWTFK-----TYFN-----EMVENKKNEIQ-EKAE 89

Qy 149 ATNYTLFTVSRNTL-----LYTLRRPVFRKMENODALVCISESVPEPIVE 195

Db 90 ATR-TGTYTCSNSGLTSSIIYVVRDPAKFLVGLPLFGK-EDSDALVRCPLDPO-VSN 146

Qy 196 WVLCDSQGESKKEESPAP-----VKKEEVHLHFGTDIRCCARNELGRECTRL--- 244

Db 147 YSLIEDGKSLPTDLTFVNPNAKAGITIKNVKRAYHRLC---VRCAAQ---RDTWLHSD 199

Qy 245 -FTIDLNQI-----PQTLPLQL--FLKVGPELWTRCAVHVNHGFLTW-----E 286

Db 200 KFTLVKREAIKAIQVSVVPETSLHLLKGGDTFTVYCTKDVSTSVNSMLKMNPPQOHAQ 259

Qy 287 LENKALEEGNPEMSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPQSALVT--- 343

Db 260 VKNSHWRGDF---NYRQETIT-----ISSARVDDSGVFCVYANTTFCGSANVTTLK 309

Qy 344 IVKGFINATN-SSDEYEDQYEEFCFSVRKAYPQ-IRCTWTFSRKSFPCEQKGLD--- 398

Db 310 VVEKGFINISPVKNITVFTVDGENVDLVVEYEAYPEHQWYIMNRT--SANKGKYVK 367

Qy 399 -NGYSISKCNKH-----QGEIFHAENDDAQTKMFTNIRKPKQVLA--EASAS 448

Db 368 SDNKSIRYVNLRLTLKGTGCTYTLVNSDASASVTFFNVYVNTKPEILTYDRLING 427

Qy 449 QASCFSDGYPLPSWTWKCKSDKSPNCTEITE-GVNNRKNRKFVGQWSSSTLNMSEAI 507

Db 428 MQCVAEGFPEPTIDWYFCGAEORCTTPSPVDVQVONVSVSPFGKLVQSSIDSVFR 487

Qy 508 KGPLVKCCAYNSLGTSCETILLNSPGPPF---IQDN-----ISFYATIGVCLLFIV 557

Db 488 HNGTVECKASNDVGKS--SAFFN---FAFKEQIOAHTLFTPLLIGVVAAGANGIIVM 541

Qy 558 LTLILCHYKKQPRYESQLQMV-OVTGSSDNEYFYVDREYEDLKWEPRENLEPFQKVL 616

Db 542 LT---YKYLQKPMYEVQWKVVEEING---NNVYIDPTQPLFDHKKWEPNRLSFQTL 594

Qy 617 GSCAFKVMNATAYGISKTGVSTQVAVKMLKEKADSSEREAALMSLKMNTOLGSHENIVN 676

Db 595 GAGAFKGVVEATAYGLTKSDAAMTVAVKMLKPSAHLTEREALMSLKLVSILGNHNNIYN 654

Qy 677 LLGACTLSGPIYILIFYCCYGDLNLYRSKRKEF-----HRTWTEIFKEHNPFFYTFQ 730

Db 655 LLGACTVGGPTLVITEYCCYGDLNLYRRKRDSEIFSKQEQAFAALYKLLHSTESCD 714

Qy 731 SHPNSSM---PGSREVOIHPDSDQISLGHNSFHSDEIEYENOKLEEEEDLNVLFTE 787

Db 715 S-SNEYMDKMPGVSYV-VPTKDKRRSARIDSY-----IERDVTPTAMEDDEL-ALDLD 766

A:Cross-references: GB:S78839; NID:g244084; PIDN:AAB21234.1; PID:g244085
A:Note: sequence extracted from NCBI backbone (NCBIN:78839, NCBI:P:78842)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: B41815
A:Molecule type: DNA
A:Residues: 637-641, "SPELPW" <SP2>
A:Cross-references: GB:S78843; NID:g244086; PIDN:AAB21235.1; PID:g244087
A:Note: sequence extracted from NCBI backbone (NCBIN:78843, NCBI:P:78844)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: C41815
A:Molecule type: DNA
A:Residues: 556-560, "GGDKWK" <SP3>
A:Cross-references: GB:S78845; NID:g244088; PIDN:AAB21236.1; PID:g244089
A:Note: sequence extracted from NCBI backbone (NCBIN:78845, NCBI:P:78846)
A:Note: disease-related mutant from patient with piebaldism
A:Giebel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.
Oncogene 7, 2207-2217, 1992
A:Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth factor receptor) proto-oncogene.
A:Reference number: 137948; MUID:93064697; PMID:1279499
A:Accession: 137948
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-976 <RES>
A:Cross-references: EMBL:X69301; NID:g34089; PIDN:CAA49159.1; PID:g825686
A:Note: an alternative splice form omitting residues 510-513 is described
R:Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.
Jpn. J. Cancer Res. 84, 1136-1144, 1993
A:Title: Characterization of the promoter region of the human c-kit proto-oncogene.
A:Reference number: 156954; MUID:94103107; PMID:7506248
A:Accession: 156954
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:S67773; NID:g459358; PIDN:AAB29529.1; PID:g459359
R:Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukui, K.
Hum. Mol. Genet. 2, 1499-1500, 1993
A:Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-oncogene.
A:Reference number: 154336; MUID:94061059; PMID:7694728
A:Accession: 154336
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 242-250 <RE3>
A:Cross-references: GB:S67686; NID:g460545; PIDN:AAD13996.1; PID:g4261696
C:Genetics:
A:Gene: GDB-KIT
A:Cross-references: GDB:120117; OMIM:164920
A:Map position: 4q12-q12
A:Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 600/1
A:Note: defects in this gene may result in piebaldism
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related
A:Title: Tyrosine-specific protein kinase
F:1-976/Product: protein-tyrosine kinase
F:1-509,514-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <M1>
F:23-976/Product: protein-tyrosine kinase kit #status predicted <M2>
F:23-520/Product: protein-tyrosine kinase kit #status predicted <EXT>
F:51-99/Domain: immunoglobulin homology <IMM1>
F:129-188/Domain: immunoglobulin homology <IMM2>
F:226-292/Domain: immunoglobulin homology <IMM3>
F:328-394/Domain: immunoglobulin homology <IMM4>
F:423-493/Domain: immunoglobulin homology <IMM5>
F:521-543/Domain: transmembrane #status predicted <TM>
F:544-976/Domain: intracellular #status predicted <INT>
F:587-931/Domain: intracellular #status predicted <KIN>
F:595-603/Domain: protein kinase ATP-binding motif
F:58-97,136-186,233-290,428-491/Disulfide bonds: #status predicted
F:130,145,283,300,320,352,367,463,486/Binding site: carbohydrate (Asn) (covalent) #s
F:623,640,792/Active site: Lys, Glu, Asp #status predicted
F:797,810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity 32.7%; Pred. No. 1.4e-53;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;
Qy 47 VCKSSYPWSESPDLGALRPOSSGTYEAAAEVDVYSASITLQVLDAPGNISCLAV 106
Db 20 VOTGSQSPVSPG-EPSPSIHPGKSDLI-----VRVGEIRLLCTDGFVK--WT 67
Qy 107 FKSSLLNCPHFQDLQNRGVVSMVILKMTETQAGEYLLFIQSEATN---YTILFTVSIRNT 163
Db 68 FE-----ILDETENKQNEWIT-EKAEATNGKYTCTCNKHGLSNS 106
Qy 164 L-----LYTLRRPYPRKMNODALVCISPEPEPIVWVLCDSQGESKEE----- 209
Db 107 IYVVRDPAPKFLVDRSLYGG-EDNLTIVRCPLTDPE-VTNYSKGCQKPLPKDLRFIP 164
Qy 210 ---SPAVVKEEKVILHFGTDIRCCARNELGRECRFLFTIDL---NOTPQTLTQ--L 260
Db 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFLKVRPAFAKVPVVSVKASY 221
Qy 261 FLKVGELWIRCKAVHVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFV 317
Db 222 LLREGSEFTVTCTIKDVSSSVYTWKRENSOTKLOEKYNSWHHGDNYERQAT---LTI 277
Qy 318 SSVARNDTGYTCSSSKHPSQSALVT---IVGKGFNA-----TNSSEYEDIOYE 365
Db 278 SSARVNDSGVFCMYANNTFGSANVTTLEVDKGFNIPFMINPTTVFVNDGENVDL--- 333
Qy 366 EFCFSVFRKAYPO-IRCTWTFSRKSF--PCEOKGLONGYSISKFNHKK-----OPGE 415
Db 334 ---IVEYEAFFKPEHQOIIYNNRTFTDKWEDYPKSENEINIRYSELHLHLRLKGTGGT 389
Qy 416 YTFHAENDDAQFTKMTLNIIRKPOVLA--EASASQASCFSDGYPPLPSMTWKCSOKSPN 473
Db 390 YTFVNSDSDVNAIAINIVYVNTKPEILTYDLRVNGMLQCVAAAGPEPTIDWYFCPTQEQ 449
Qy 474 CTBEITE-GVNNRKANKRVFGQWSSSTLNMSPAIKGLVKCCAYNSLGTSCETILLNSP 532
Db 450 CSASVLPVDVQTLNNSGPPGKLVQSSIDSSAFKNGTVECKRAYNDVGKT--SAYEN-- 505
Qy 533 GPFPFTODN-----ISFYATIGVCLLFTVLVTLTLLCHKYKQFRYESOLOM 578
Db 506 --FAFGNNKEQIHPHTLFTPLDIGFVIVAGMCIIVMILT---YKLOKPMFEVQWKV 559
Qy 579 V-OVTGSSDNEYFYDFREYEDLKWEFFRENLEFGVLGSGAFKVMNATAYGISKTGV 637
Db 560 VEING---NNYVYIDPTQLPYDHKEFFPNRLSFGKTLGAGAFKVVEATAYGLIKSDA 616
Qy 638 SIQVAVKMLKEKADSGEREALMSLKMNTOLGSHENIVNLLGACTLSGPYILFIEYCCYG 697
Db 617 AMTVAVKMLKPSAHLTEREALMSLKVLSYLGNNHNVNLLGACTIGGPTLVITEYCCYG 676
Qy 698 DLLNLYLRKREKFRHWTE-----IPKEHNFSPYPTFQSHPNSSM---PGSREVOIHPD 748
Db 677 DLLNFLRRKRDSTICSKQEDHAEALYKLLHLSKESCDSTNEYDMKMGVSVV-VPTK 735
Qy 749 SDQISGLHNSHSEDEIEYENOKRLEEDLNLTFTEDLLCFAYQVAKGMFELEKSCV 808
Db 736 ADRRRSVRIGSY-----IERDVTPTAIMEDEL-ALDLELLSFSYQVAKGMFLASRNCI 789
Qy 809 HRDLAARNVLVTHGVVVICDFGLARDIMSDSNVYVVRGNARLPVKKMAPESLFGIYTIK 868
Db 790 HRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVVRGNARLPVKKMAPESIFNCVTFE 849
Qy 869 SDVMSYGIILWEIPLVSLGVNYPGIPDANFYKLLQNGFKMDQPFYATEEYIIIMQSWAF 928
Db 850 SDVMSYGIILWEIPLVSLGVNYPGIPDANFYKLLQNGFKMDQPFYATEEYIIIMQSWAF 928
Qy 929 DSRKRPSFPNLTSLGCOLADAEAMVQNV 958
Db 910 DPLKRPTKQIVQVIEKQISESTNHIYSNL 939

RESULT 8

151703
C-kit-related kinase 1 (Xkrk1) - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I51703
R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.
Mech. Dev. 50, 217-228, 1995
A:Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem cells
A:Reference number: I51703; MUID:95344996; PMID:7619732
A:Accession: I51703
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-954 <BAK>
A:Cross-references: EMBL:248770; NID:g763033; PIDN:CAA8688.1; PID:g763034
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology
C:Keywords: ATP
F:573-915/Domain: protein kinase homology <KIN>
F:583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1226; DB 2; Length 954;
Best Local Similarity 33.1%; Pred. No. 2.2e-52;
Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;

QY 107 FKHSLLNCPHFQDLQNRGVYS-----MVLKMTETQAGEYLLFIQSEATNYTLFTVSI 160
DQ 107 FKHSLLNCPHFQDLQNRGVYS-----MVLKMTETQAGEYLLFIQSEATNYTLFTVSI 160
DQ 52 FOKSGLMKKPR-DLKSRLNNSETDOFFVIKADLRHIGRYICTNTQENTSV--SLFV 108
QY 161 RNTLYTLRRPFRKMNODALVCISVPEPIVWLCDSQGESCKE-----ESPA 212
DQ 161 RNTLYTLRRPFRKMNODALVCISVPEPIVWLCDSQGESCKE-----ESPA 212
DQ 109 KDPARPFLDIPEDIVTEGADTGMCFPTDPMIDIAEKCD--GSLPDPNFTTDTIEAGI 166
QY 213 VYKKEKVLHFLGFDIRCCARNELGR-ECTRLFTIDLNQTPOTTLPLQLF-----KV 264
DQ 213 VYKKEKVLHFLGFDIRCCARNELGR-ECTRLFTIDLNQTPOTTLPLQLF-----KV 264
DQ 167 TIKTVOLAPDSY-----VCSNKSQTVKSKSTFSIHVAPVVK-KVPTVFLSKSRQLVK 220
QY 265 GEPFLIRKAVHNVHFGGLTW-ELENKALEEGNYFMSTYSYNTMIRILFAFVSSVARN 323
DQ 221 GEPFVTCVAVLDFSTVKAQDLVREGVTKQANFRSNVFSYNLTLS-----DGVFYS 274
QY 324 DRGYTCSSKHPQS---SALVTIVGKFINAT-NSEDEYIDQYEEFCFVFRKAYPQ- 378
DQ 275 ESRTTQCAENAIQGVNATFTLDIVGVNLTVENTTISVNAQDNLVKLVYDAYPHP 334
QY 379 IRCTWTESKSPCEQKGLDNGYSISK-----FCNHKH-----QPEYIPHAENDDA 425
DQ 335 DGVWTFYFNETL-----LNTSDHYATKDEGNRIYVSELHLIRLKGKGYVTFYTTNSDD 390
QY 426 QFTKFTLNRKPKQVLAESASQAS--CFSDGYPLPSWTWKCKSDKSPNCTEETEGVM 483
DQ 391 DASVFNIOVKTRPELIIAERTSEGTLOCVATGFPVPAIQWFCPGSEQRCTDYPPLSPV 450
QY 484 NKK--ANKRVQGWTSSTLNNSEAIKGLVKKCAVNSIGTSCETILLNSPGFPF-IO 540
DQ 451 NEKFIQENSLSGRIVVESTIDVNDLKNKTGQCVASNEVESYV-----FSEAIKE 502
QY 541 N-----ISFVATIGVCLLFTVLTLLCHIKYKKQFYESOLOMV-QVTGSSDNE 588
DQ 503 KURTHLFTPLLIGFTAAAGLMCIAVAVL-----MYKLOKPYEIQWKVVEEING---NN 555
QY 589 YFYVDFREYEDLKWFFRENLEFGKVLGSAFGKVMNATAYGISKTGVSIOVAVKMLKE 648
DQ 556 YVYIDPTQLPYDNKWEFFRDRLCFGKILGAGAGKVEATATAGLLKDESLRIVAVKMLKP 615
QY 649 KADSSEREAALMSELKMMTOLGSHENIVNLGACTLSGGTYLIFYECCYGDLLNLYRSKRE 708
DQ 616 SAHSTEREALMSELKVLVSYLGHKHNIVNLGACTVGGPLTVITEYCCYGDLLNLYRSKRD 675
QY 709 KPHRTWTEIFKEHNSFYPTTFOSHNS-----SMPGSEVQIHPDSQI-----752
DQ 676 SP-----ICPKFDNSEAALYKKNLLNTRDMGCEGMSYIDMKPAVSVVWPT 721
QY 753 --SGLHGNFSHSEDETEYENQKRLEEEEDLNVLTFDLLCFAYQVAKGMEFFLEKSCVHR 810

Db 722 KTKRRSGSFGDQDV-----SVSIPEDDL-ALDTEDLNFSYQVAGMNFASKNCIHR 775
QY 811 DLARNVLVTHGVKVKICDFGLARDIMSDSNVYVGRNARLPVKWMAPESEFEGYITKSD 870
Db 776 DLARNILLTHGRITKICDFGLARDIRNDSNVYVKGARLPVKWMAPESEFHCYTFESD 835
QY 871 VWSYGILLWEIFSLGVNPGIPVDANFYKLQNGFKMDQPFYATEEYIIMQSCWAFDS 930
Db 836 VWSYGILLWEIFSLGVNPGIPVDANFYKLQNGFKMDQPFYATEEYIIMQSCWAFDS 930
QY 931 RKRSPFNLTSTFLGCOLADAE 951
Db 896 LKRPTKQIVQMVEQQLSDSK 916

RESULT 9
TVCTMD
macrophage colony-stimulating factor 1 receptor precursor - cat
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 13-Jun-1997
C:Accession: A31636
R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.
Cell 55, 965-977, 1988
A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are required
A:Reference number: A31636; MUID:89077553; PMID:2849512
A:Accession: A31636
A:Molecule type: mRNA
A:Residues: 1-980 <WOO>
A:Cross-references: EMBL:X03663
C:Genetics:
A:Gene: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming prote
fic protein kinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <
F:24-309/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-484/Domain: immunoglobulin homology <IMM5>
F:510-534/Domain: transmembrane #status predicted <TM>
F:535-980/Domain: intracellular #status predicted <INT>
F:577-915/Domain: protein kinase homology <KIN>
F:585-593/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted
F:45-73,94,153,275,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #s
F:613,630,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1225.5; DB 1; Length 980;
Best Local Similarity 33.5%; Pred. No. 2.4e-52;
Matches 321; Conservative 168; Mismatches 321; Indels 149; Gaps 34;

QY 83 VDVSASITLQVLVDAPGNISCLW---VFKHSSLNCP-----HFDLQNRGVVSMVIL 131
Db 32 VEPGTVTLRCV---GNGSVENDGPISPHNLDLDPSSILITNNATFQNTGYHCT-- 85
QY 132 KMTETQAGEYL--LFTQSEATNYTIL---FTVSRINLTLLRPPYRKMNQDALV-CI 185
Db 86 EPGNPOGGNATHLYKDKPARPKVLAQEVTV-----LEGQDALLPCL 128
QY 186 SESVPPPIVE--VWLCDQSQESCKE-----SP---AVVKEEKVLHFLGFDIRCCARN 235
Db 129 ---LTPALEAGVSLVRVGRPVLRQTNYSFSPWHGTHHKAFIENHVY---QCSARV 181
QY 236 ELGRECTRL-----FTIDLNQTPOTTL-PQLFLVK-GEPLWIRCKAVHNVHFGTLWELE 288
Db 182 D-GRTVTSMGILWKVKDLSGPAFLLEPAELVRIQGEAAQIVCSASNIDVNFV-----235

Db 560 SYTFIDPTQLPYNKWEFFPRNNLOFGKTLGAGAFKGVVETAFGLGKEDAVLKVAVKMLK 619
Qy 648 EKADSSREALMSKMMTOIGSHENIVNLGACTLSGPIYLIFECYCYGDLNLYLSRKR 707
Db 620 STAHADKEALMSKIMSHLQGHENIVNLGACTHGGPVLVITCYCYGDLNLYLSRKA 679
Qy 708 E-----KFWRTWTEIFKEHNFSEFYPTFQSHPNSSMPCGSRVQIHP 747
Db 680 EAMLGPSLSPQDPEGVDYKNIHLKKYVRDSDGFS-----SOGVDYV 724
Qy 748 DSDOISGLHGNFSHSEIEYENOKRLEEDLNVLTFEDLLCFAYQVAKGMFLEKSC 807
Db 725 EMRPVSTSSDSF-SEQDLKEDGRPLE-----LRDLLHFSQVAQMAFLASKN 774
Qy 808 VHRDLAARNVLTHTGKVVYKICDFGLARDIMSDSNVYVRGNARLPVKMMAPESLFEIYI 867
Db 775 IHRDVAARNVLTHTGKVVYKICDFGLARDIMSDSNVYVRGNARLPVKMMAPESLFEIYI 834
Qy 868 KSDVWSYGIILLWEIFSLGVPNPYGPIDVANTYKLIQNGFKMDQPFYATEEYIIMQSCWA 927
Db 835 QSDVWSYGIILLWEIFSLGVPNPYGPIDVANTYKLIQNGFKMDQPFYATEEYIIMQSCWA 894
Qy 928 FDSKRKSPNLTSLFCOL-ADAEEMQYQV 958
Db 895 LEPTHTPTFOQICSLQEQAOEDRRERYTNL 926

RESULT 12
JN0677
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken
N;Alternate names: tyrosine kinase receptor kit
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: JN0677
R;Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.
Gene 128, 257-261, 1993
A;Title: Cloning and expression of the chicken c-kit proto-oncogene.
A;Reference number: JN0677; MUID:9329295; PMID:7685729
A;Accession: JN0677
A;Molecule type: mRNA
A;Residues: 1-960 <SAS>
A;Cross-references: DBJ:DJ1225; NID:g303532; PIDN:BAA02506.1; PID:g303533
A;Experimental source: brain
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: Atp; autophosphorylation; glycoprotein; kinase-related transforming protein;
roten kinase
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-960/Product: tyrosine kinase receptor #status predicted <MAT>
F;314-380/Domain: immunoglobulin homology <IMM>
F;573-916/Domain: protein kinase homology <KIN>
F;581-589/Region: protein kinase Atp-binding motif
F;76,135,149,269,286,306,318,338,356,453,469/Binding site: carbohydrate (Asn) (covalent)

Query Match 23.0%; Score 1215; DB 1; Length 960;
Best Local Similarity 32.3%; Pred. No. 7.5e-52;
Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;

Qy 66 ALRPOSSCTVYEAAREVDVSAITLQVLVDAPGNISCLWVFKHSS-----LNCQPHFD 119
Db 18 SLIPAGVSPVHESSLVNKGGEELKCNKEGP-----VTWNFQSDPSAKTRISNEKEWH 73
Qy 120 LQNRGV--VSMVILKWTETQAGEYLLFTQSEATNTILFTVSIRNTLLTLRRPFRKME 177
Db 74 TKNATIRIGRVECKSGKSIVNSFYVFKDP---NVLFLV---DSLIV-----GKED 119
Qy 178 NODALVCISVPEP-IVEWVLCDQSGECKSESPAVKKEE----KVLHELFGTDIRCC 232
Db 120 SDILLVC---PLTDPDLNFTLRKCDGKPLPKNMTFIPNPQKGIINKVQSRFKGCYQCL 176
Qy 233 AR-NELGRECTRLFLIDLNQT-PQTLLPOL-----FLKVGEPWLIRCKAVVNHGFL 283
Db 177 AKHNGVEKISEHIF---LNVRPVHALPVTILSKSYELLKGEFEFVTCIITDVS SVKA 233

Qy 284 TWELNKALEBGNFYFEMSTYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHP--SQSAL 341
Db 234 SWISYSAIVTSKRNUGDGYERK---LTNIRKSVGNDSGEFTC-QAENPFKTKNAT 288
Qy 342 VTI--VGKGFIN--ATNSSSEYEDIDQYEEFCFSYFRKAYQIR-CTWTFSRKSPCEOK 395
Db 289 VTLKALAGFVRLFATMTTIDINAGONG--LTVEYEAYPKPKKEEVMMYMETL---QN 343
Qy 396 GLDNGYSTSKFCNKH-----OPGEYIIFHAENDDAQFTKMTLNIRRPQVLA 443
Db 344 SSDHVYKFKETVGNNSYTSSELHLRLKGTGGIYTFEVSNSDASSSVTFNVYVTKPEILT 403
Qy 444 EASASQ--ASCFSQGYPLPSMTWKCKSDKSPNCTEITEGVNMRKANRKFVFGQWSSSTL 501
Db 404 LDMGNDILQCVATGFPAPTIYWFPCGTEQRCLOSTPISPMVKVS-----YTNSSVP 457
Qy 502 NMSEAIKGLVKCCAYNSLGTSCETILLNSPGFPF---IQDNISFYA-----TIGV 550
Db 458 SPERILVESTVNASMFKSTGTICCEASSNGDKSSVFENFAKEQIRHTLFTPLLIAFGV 517
Qy 551 CLLFTVLTLLICHYKKOFYESOLQMV-QVTGSSDNEYFVDFREYEDLKWEPREN 609
Db 518 AAGLCIIIVMILVYIYLOKPKYEVQKVVVEING---NNYVYIDPTQLPYDHKWEFPRNR 574
Qy 610 LEFGVLGSGAFKVMNATAYGISKTGYSIQVAVKMLKEKADSSREALMSLKMNTOLG 669
Db 575 LSFGLTLAGAFKGVVEATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSYLG 634
Qy 670 SHENIVNLGACTLSGPIYLIFECYCYGDLNLYLSRKRKF---HRTWTEIFKEHNFSE 725
Db 635 NHINIVNLGACTIGPTLVITEYCYCYGDLNLYLSRKRDSFCPKHEEHAAEAVENL-- 692
Qy 726 YPTFQSHPN-----SMPSGREVOIHPDSQISGLHGNFSHSEDEIEYENOKRLEEE 778
Db 693 --LHQAEPTADAVNEYMDMKPGSVAVPPKADKRPVKSQSYTDQD---VTLMSLEDDE 746
Qy 779 DLNVLTFEDLLCFAYQVAKGMFLEKSCVHRDLAARNVLYTHGKVVYKICDFGLARDIM 838
Db 747 --LALDVEDLLSFSYQVAKGMSFLASKNCHRDLAARNILLTHGRITKICDFGLARDIRN 804
Qy 839 DSNVYVRGNARLPVKMMAPESLFEIYITKSDVWSYGIILLWEIFSLGVPNPYGPIDVAND 898
Db 805 DSNVYVRGNARLPVKMMAPESLFEIYITKSDVWSYGIILLWEIFSLGVPNPYGPIDVAND 864
Qy 899 YKLIQNGFKMDQPFYATEEYIIMQSCWAFDSKRKPSNLTSLFCQCLADAEEMQYQV 958
Db 865 YKMIKREGYRMFSESPPEMYDIMKSCWDADPLQRTFKQIVQLIEQQLSDNAPRVAN- 923
Qy 959 DGRVSECPHTYON 971
Db 924 ---FSTPPSTOQN 933

RESULT 13
T30816
macrophage colony-stimulating factor receptor - Japanese pufferfish
C;Species: Fugu rubripes (Japanese pufferfish)
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
C;Accession: T30816
R;How, G.F.; Venkatesh, B.; Brenner, S.
Genome Res. 6, 1185-1191, 1996
A;Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for
A;Reference number: 220882; MUID:97129405; PMID:8973913
A;Accession: T30816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-975 <HOW>
C;Cross-references: EMBL:U63926; NID:g1752706; PID:g1752708; PIDN:AAC60063.1
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

Query Match 23.0%; Score 1213; DB 2; Length 975;
Best Local Similarity 31.7%; Pred. No. 9.6e-52;

Matches 325; Conservative 177; Mismatches 333; Indels 190; Gaps 37;

```
QY 33 IKCVLINHKNNDSSVGKSSYP---MVSESPEDLGC--ALRP--QSSCT---VVEAAAV 81
Db 47 LKC-----GGDPVNMOTRLPKHKRYMSRSPGNLRTIRVAPRTAEFTGYTKCFYSAAQ 100
QY 82 EVDVSASITLQVLVDAPGNISCLWFKHSSLNCPHFDLQNRGVSMVLKMTETQAGFY 141
Db 101 HRLTSSV--HVYKDPNRV--FW-----TSSTSLVRVRKEGDEY 136
QY 142 L--LFIOSEATNYTLTIVSIRNTLLYLRLRPYFRKMENQDALVCISESRPEPIVWVL 198
Db 137 LPLCLLTDPATDGL-----RMDN-----GTTVP-PEMNYTV 168
QY 199 CDSQESCKEESPAVVKKEKVLHFGTDIRCCARNELGRECTRLFTDLNQTPTLLP 258
Db 169 YRHGILIRLOPS-----FNADYVCTAKVGKVEKTSFTFSINVLQIRFP-P 215
QY 259 QLFK-----VGEPLWIRKAVHNVHGFGLTWELNKA---LEEGNYFEMSTYSTRT 308
Db 216 YVFLMEDEYVRIVGEELQIRCMTHNPENYNVNTYTKSRVTIEE---RVRSSGENRL 271
QY 309 MIRILFAVSSVARNDTGYTCSSSKHP---SQSALVTIVGKFI-----NATNSSE 357
Db 272 DIQSILT-ISAVDLADTGNISCIGTNEAGVNSNTYLLVWEKPYIRLWPQLIPKLASOQL 330
QY 358 DYEIDQVEEFCFSVRFRKAYPOIR-CTWTFSRKSPCEOKGLDNGYSISKFCNHKHPGEY 416
Db 331 SVEVNEGDELGVWVEAYPQIDTHRWHTPTSPSTSMGHEIYHARLQKRMNAQEQ-QGY 389
QY 417 IFHAENDDAQFTKMTLNIIRKP-QVLAESAASQSCFSDGYPFSPWTWKKCDKSPNCT 475
Db 390 TFYAKSNLANGSISFHVKMYOKPIAVRWENITLTCTSFQYPAPIIWOCSGIRPTCN 449
QY 476 EEITEGVNRR-----KANRKFQGWVSSSTLMNSAIFGLVKCAYNLSLSCETI 527
Db 450 GNNT-GLPKQNHQALTVQVOREEYGAVEVESFTVGLSNHRMTVECFANLVGVSSDTF 508
QY 528 LLNSPGPFPTODNISFVATI---GVCLLFIVTLTLLCHYKKQRYESQLOLWQVYTG 583
Db 509 TVE-----VSDKL-FTSLIGAGVLAIFELLLVFL--TKYOKRFEIRWKIE--A 557
QY 584 SSDNEYFYVDFREYEDLKWEPFRENLEFGVLGSGAFKVMNATAYGISK-TGVSIQVA 642
Db 558 REGNNTFIDPTQLPYNEKEWEPFDKLGKLVGAGAFGKVVEATAFGLGEDKONTLRA 617
QY 643 VKMLKEKADSEREALMSLKMWTOLGSHENIVNLGACTLSGPIYLIYFECYCGDLNY 702
Db 618 VKMLKANAHSDEREALMSLKLHLGHQHNIVNLGACTYGGPVLVITEYCSLGLDLNF 677
QY 703 LRSKREKFRHTWTEIFREHNFSPYPTQSHPN-----SSMPSREVO 744
Db 678 LRQAEFTVNLVMI-----PEIMENSDYKNICNQKWIRSDSGISSTSSSTYLE 728
QY 745 IHPDSQISGLHGNSFHSDEIEYENOKRLEEDLNVLTFEDLLCFAYQVAKGMFELEF 804
Db 729 MRPS-----QQSHIEASGRKSLCEDNGDWPLDIDLLRFSLOVAQGLDFLAS 775
QY 805 KSCVHRDLARNVLVTHGKVVVKICDFGLARDIMSDSNVYVVRGNARLPVKWMAPESLFEG 864
Db 776 RNCIHRDVAARNVLLTKRVAKICDFGLARDIMSDSNVYVVRGNARLPVKWMAPESIDCV 835
QY 865 YTIKSDVMSYGILLWEIFSLGVNYPGIPVDANYFKLQNGFKMDQDPYATEEYIIMQS 924
Db 836 YTVOSDVMMSYGILLWEIFSLGVNYPGIPVDANYFKLQNGFKMDQDPYATEEYIIMQS 895
QY 925 CWAFFDSKRKPSFNLTSFLGCOLA---DAEAMYNQV-----DGRVSECPTHYQNRPE 975
Db 896 CWNLEPTERTFMSISQMINRLLGGQDEQEKLIYRNVOPEQVAGEACDEPKRYD--PPC 953
QY 976 SREMD 980
Db 954 ERSDD 958
```

RESULT 14

S16385

macrophage colony-stimulating factor 1 receptor precursor - rat

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSF-1R

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Jun-2000

C:Accession: I60321; S16385

R:Borycki, A.G.; Guiller, M.; Leibovitch, M.P.; Leibovitch, S.A.

Growth Factors 6, 209-218, 1992

A:Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis an

A:Reference number: I60321; MUID:93001225; PMID:1389227

A:Accession: I60321

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-978 <RES>

A:Cross-references: EMBL:X61479; NID:g57543; PID:CAA43706.1; PID:g57544

A:Note: in Genbank entry RRCSE1, release 113.0, the source is designated as Rattus ra

A:Note: submitted to the EMBL Data Library, August 1991

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-re

protein; tyrosine-specific protein kinase

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <

F:35-86/Domain: extracellular #status predicted <EXT>

F:120-179/Domain: immunoglobulin homology <IMM1>

F:217-280/Domain: immunoglobulin homology <IMM2>

F:316-381/Domain: immunoglobulin homology <IMM3>

F:410-485/Domain: immunoglobulin homology <IMM4>

F:516-535/Domain: immunoglobulin homology <IMM5>

F:536-978/Domain: transmembrane #status predicted <TM>

F:578-978/Domain: intracellular #status predicted <INT>

F:586-978/Domain: protein kinase homology <KIN>

F:42-84/127-177,224-278,417-483/Disulfide bonds: #status predicted

F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #stat

F:614,631,776/Active site: Lys, Glu, Asp #status predicted

F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.6%; Score 1194; DB 2; Length 978;

Best Local Similarity 33.1%; Pred. No. 8e-51;

Matches 321; Conservative 154; Mismatches 332; Indels 164; Gaps 32;

QY 83 VDVSASITLQVL---VDAPGNISCLWFKHSSLNCPHFDLQNRGV-----SWILKM 133

Db 32 VEPGETVTLRCVNSGSEWMDGPISPYWTLDPES---PGSTLTTRNATEFKNTGYRCTEL 87

QY 134 TETOAGEYL--LFIQSEATNYTL---FTVSIRNTLLYLRLRPYFRKMENQDALVCISES 188

Db 88 EDPMAGSTTIHLYVKDPAHSMWLLAQEVTV-----VEGQEA----- 124

QY 189 VPEPIVWLCDQSQESCKEESPAVVKK-----EEKVLHFLFGTDIRCCAR 234

Db 125 LPLCLTDPALKDSV-SLMREGGRQVLRKTVYFFSAWRGFIIRKAKVL----DSNTVCKT 179

QY 235 NELGRECTRL-FTIDLN---OTPOTTL-PQLFLKY-GEPLWIRKAVHNVHGFGLTWEL 287

Db 180 MVNGRESTSTGWLKVRNVHPPEPQIKLEPSKLVIRGEAAQIVCSATNAEVGFNVILKR 239

QY 288 ENKALE-----EGNYFEMSTYSTNTRMIRILFAFVSSVARNDTGYTC-SSSKHPSQS 339

Db 240 GDTKLEIPLNSDFQDNYK-----KVRALSLNAVDFQDAGIYSCVANDVGRT 288

QY 340 ALVT--IVGKGFINATN--SSEDEYIDQVEEFCFSVRFRKAYPOIR-CTWTFSRKSPCEOK 395

Db 289 ATMFQVVEAYNLTSQSLLOEVSVDLSLTVHADAYPSIQHYNNWYTLGPFPE-DQR 347

QY 396 GLD-----NGYSISKFCNH--KHQPGEYTFHAENDDAQFTKMTLNIIRKRPQVLAES 446

Db 348 KLEFITORAIRYTFKFLNLRVKASEAGYFLMAQNAKGNWNLITELTLRYPEVSVTHM 407

QY 447 ASQAS-----CFSDGYPLPSWTWKKCDKSPNCTEITEGVVN-----RKANRKFVGQWSS 498

Db 408 PVNGSDVLCFQVSGYPPQSVTWMCRGHTDRCDQAALQVNDTHPEVLSQKPPDKVLIQ 467
Qy 499 STLNMSEAIGFLVKCCAYNSLGTSCFPIILNSPGPPFFIQDNISFYATIIGVC---LLFI 555
Db 468 SOLPIGTLKHNMTYFCFKTHNSVGNSSQYFRAVSLGSKQOLPDESFTFPVWVACSVMSLL 527
Qy 556 VVLLILLCHYKKOFRYESQLOMVOVQVSSDNEYFYVDREYEDLKWEPFRENLEFCVKV 615
Db 528 VLLLLLLLYKKPKQYQVRWKIIE--RYEGNSYTFIDPTQLPYNEKWEFFPRNNLOFKGT 585
Qy 616 LGSAGFGVMNATAYGISKTVSIOVAVKMLKEKADSSERALMSSELKMMTQOLGSHENIV 675
Db 586 LGAGAGKVVEATAFGLCKEDAVLKVAVKMLKSTAHADKEALMSSELKIMSHLGOHENIV 645
Qy 676 NLLGACTLSGPIYLIIFYECCYGDLLNYLRSKRE-----KPHRTWT 715
Db 646 NLLGACTHGGPVLVITEYCCYGDLLNLRRAEAMLGPSLSPGQSDSGSYKNTHLEKK 705
Qy 716 EIFKEHNFSPYTFQSHNSNMPGSRVQIHPDSQIISGLHCNSFSEDEIYENOKRLE 775
Db 706 YVRDSGFS-----SQGVDTVVEMRPVSTSSDSFFKQD-LDKEPSRPLE 749
Qy 776 EEDNLVTFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARD 835
Db 750 -----LWDLHFSSQVAGGMAFLASKNCIHRDVARNVLLTSGHVAKIGDFGLARD 800
Qy 836 IMDSNNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVD 895
Db 801 IMDSNNYVVRGNARLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPGILVN 860
Qy 896 ANFYKILONGFKMDOPFYATEEIIIMQSCWAFDSKRPRSPFNLTSLFCCQLADAEAM 955
Db 861 NKFYKLVKDGVOYMAQVFPAPKNIIYSIMQSCWDLPTRRPTFOQICPLL-----QEQARL 914
Qy 956 ONVDGRVSECP 966
Db 915 ERRDQYANLP 925

RESULT 15

TVMSND

Macrophage colony-stimulating factor 1 receptor precursor - mouse

N:Contains: protein-tyrosine kinase. (EC 2.7.1.112) csflr/fms

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Jun-2000

C:Accession: S01880

R:Rothwell, V.M.; Rohrschneider, L.R.

Oncogene Res. 1, 311-324, 1987

A:Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.

A:Reference number: S01880; MUID:88217329; PMID:2966922

A:Accession: S01880

A:Molecule type: mRNA

A:Residues: 1-976 <ROT>

A:Cross-references: EMBL:X06368

C:Genetics:

A:Gene: fms

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;

fic protein kinase

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted <NAT

F:20-515/Domain: extracellular #status predicted <EXT>

F:35-86/Domain: immunoglobulin homology <IMM1>

F:120-179/Domain: immunoglobulin homology <IMM2>

F:217-280/Domain: immunoglobulin homology <IMM3>

F:316-381/Domain: immunoglobulin homology <IMM4>

F:410-485/Domain: immunoglobulin homology <IMM5>

F:516-535/Domain: transmembrane #status predicted <TM>

F:536-976/Domain: intracellular #status predicted <INT>

F:578-914/Domain: protein kinase homology <KIN>

F:586-594/Region: protein kinase ATP-binding motif

F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted

F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #stat
F:614,631,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.4%; Score 1183.5; DB 1; Length 976;

Best Local Similarity 33.2%; Pred. No. 2.6e-50;

Matches 318; Conservative 154; Mismatches 347; Indels 139; Gaps 30;

Qy 83 VDVSASITLQVL----VDAPGNISCLWVF---KHSSLNCOPIHFDLQNRGVSVMLIKMTE 135
Db 32 VEPGETVTLRCVNSGVSEVMDGPISPIWTLDPESPSTLTTSNATFNKTYGTYRCTLEDDPM 91
Qy 136 TOAGEYLLFIOSEATNYTIL---FTVSIRNTLLYLRPRYFRKMNQDALVCISESVPP 192
Db 92 AGSTIHLVVRDPAHSNLLAQEVTV-----VEGOEAV-----LPCL 128
Qy 193 IVEVYLCDSQGSCKEESPAVVKKEKVLHELFGTDIR-----CCARNEIGRECT 242
Db 129 ITDPALKDSV-SLMREGGRQVLRKTVYFFSPWRGSIIRKAKVLDNSTYVCKTMVNGREST 187
Qy 243 RL-FTIDLN---QTPQITL-PQLFLKV-GEPLWIRCKAVHYNHGFGLTWELNKALE-- 293
Db 188 STGIWLKVRNVRPEPPQIKLEPSKLVRINGEAAQIVCSATNAEVEGFNVLLKRGDTKLEIP 247
Qy 294 -----EGNVFEMSTYSTNTRMIRILFAVSSVARNDTGYTC--SSSKHPSQSALVT--IV 345
Db 248 LNSDFQDNYYK-----KVRALSINAVDFDAGIYSCVASNDVGTTRATMNFQVV 296
Qy 346 KGKFINATN--SSEYDEIDQYEEFCFSVRKAYPQIR-CWTFSRKSFPCEQKGLD----- 398
Db 297 ESAYLNLTSQSLQLEQSVGDSLLITVHADAVPSIOHYNWYLVGPFFE-DQRKLEFITOR 355
Qy 399 --NGYSISKFCNH--KHQGEYIFHAENDDAOFTKMTINIRKPOVLAESAASQAS--- 451
Db 356 AIYRYTFKFLNLRVAKASEAGQYFLMAQNKAGNNLTFTLTLPPEVSVTWMPVNGSDVL 415
Qy 452 -CFSDGYPLPSTWTKKCDKSPNCTEETIEGVN-----RKANKRVGOWSVSSSTLNMSEA 506
Db 416 FCDVSGYPOPSVTWMECRGHTDRCDQAALHLNWDTHPEVLSQKPPDKVLIQSLPIGLP 475
Qy 507 IKGFLVKCCAYNSLGTSCETIILNSPGPPFFIQDNISFYATIIGVC---LLFIIVVLLIIC 563
Db 476 KHNMTYFCKTHNSVGNSSQYFRAVSLGSKQOLPDESFTFPVWVACSVMSLLVLLLLLL 535
Qy 564 HKYKQFRYESQLOMVOVQVSSDNEYFYVDREYEDLKWEPFRENLEKVLGSAFGK 623
Db 536 YKYKQPKYQVRWKIIE--RYEGNSYTFIDPTQLPYNEKWEFPRNNLOFKGTILGAGAFGK 593
Qy 624 VMNATAYGSKTGVSTOVAVKMLKEKADSEREALMSSELKMMTOLGSHENIVNLLGACTL 683
Db 594 VVEATAFGLCKEDAVLKVAVKMLKSTAHADKEALMSSELKIMSHLGOHENIVNLLGACTH 653
Qy 684 SGPIYLIFEYCCYGDLLNYLRSKREKFRHTWTEIFKEHNFSPYTFQSHPNSSMP----- 738
Db 654 GGPVLVITEYCCYGDHLNLRRAEAMHGP-----SLSPGQSDSGSYKNTHLE 703
Qy 739 -----GSEVOIHDPDSQISGLHCNSFSEDEIYENOKRLEEBEDLNLVTFEDL 788
Db 704 KKYVRRDSGFSQGVDTYVEMRPVSTSSDSFFKQD-LDKEHSRPLE-----LWDL 753
Qy 789 LCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDSNVYVRGNA 848
Db 754 LHFSSQVAGGMAFLASKNCIHRDVARNVLLTSGHVAKIGDFGLARDIMSDSNVYVRGNA 813
Qy 849 RLPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKILONGFKM 908
Db 814 -LPVKWMAPELSEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKILONGFKM 872
Qy 909 DOPFYATEEIIIMQSCWAFDSKRPRSPFNLTSLFCCQLADAEAMYNQVNDGRVSECP 966
Db 873 AQPVFAPKNIIYSIMQSCWDLPTRRPTFOQICPLL-----QEQARLERRDQYANLP 924

Wed May 28 09:15:00 2003

Search completed: May 27, 2003, 14:35:32
Job time : 37.5134 secs

us-09-919-408-4.rpr

Page 13

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:49 ; Search time 18.0091 seconds

(without alignments)
2286.959 Million cell updates/sec

Title: US-09-919-408-4
Perfect score: 5274
Sequence: 1 MPALDAGTVPLLVFSAM.....PFSREMDLGLLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	1 FLT3_HUMAN	P36888 homo sapien
2	4429.5	84.0	992	1 FLT3_MOUSE	Q00342 mus musculus
3	1295.5	24.6	978	1 KIT_CAPI	Q28317 capra hircu
4	1286	24.4	977	1 KIT_BOVIN	P43481 bos taurus
5	1266	24.0	975	1 KIT_MOUSE	P05532 mus musculus
6	1251	23.7	976	1 KIT_HUMAN	P10721 homo sapien
7	1247.5	23.7	975	1 KIT_CANFA	O97799 canis famil
8	1225.5	23.2	980	1 KFS_FELCA	P13369 felis silve
9	1224.5	23.2	978	1 KIT_FELCA	Q28889 felis silve
10	1224	23.2	978	1 KFS_FSVMD	P00545 feline sarc
11	1216.5	23.1	972	1 KFS_HUMAN	P07333 homo sapien
12	1215	23.0	960	1 KIT_CHICK	Q08156 gallus gall
13	1204	22.8	977	1 KFS_MOUSE	P09581 mus musculus
14	1194	22.6	978	1 KFS_RAT	Q00495 rattus norv
15	1176	22.3	1088	1 PGDS_RAT	P20786 rattus norv
16	1169	22.2	1089	1 PGDS_MOUSE	P26618 mus musculus
17	1166	22.1	1089	1 PGDS_HUMAN	P16234 homo sapien
18	1128	21.4	1087	1 PGDS_XENLA	P26619 xenopus lae
19	1105.5	21.0	1098	1 PGDR_MOUSE	P05622 mus musculus
20	1086.5	20.6	1106	1 PGDR_HUMAN	P09619 homo sapien
21	990	18.8	1338	1 KIT_FSVHZ	P17948 h vascular
22	989.5	18.8	370	1 KIT_FSVHZ	P04048 feline sarc
23	959.5	18.2	1333	1 VGR1_MOUSE	P35969 mus musculus
24	956	18.1	1336	1 VGR1_RAT	P53767 rattus norv
25	951.5	18.0	1356	1 VGR2_HUMAN	P35968 homo sapien
26	935	17.7	1348	1 VGR2_COTJA	P52583 coturnix co
27	923	17.5	1367	1 VGR2_MOUSE	P35918 mus musculus
28	917.5	17.4	1343	1 VGR2_RAT	O08775 rattus norv
29	907.5	17.2	1363	1 VGR3_MOUSE	P35917 mus musculus
30	904.5	17.2	1298	1 VGR3_HUMAN	P35916 homo sapien
31	874.5	16.6	823	1 CEK2_CHICK	P18461 gallus gall
32	849	16.1	821	1 FGR2_HUMAN	P21802 homo sapien
33	845	16.0	813	1 FGR2_XENLA	Q03364 xenopus lae

ALIGNMENTS

RESULT 1

ID	FLT3_HUMAN	STANDARD;	PRT;	993 AA.
AC	P36888; Q13414;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).			
DE	receptor FLT3 (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).			
GN	FLT3 OR STK1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-94119906; PubMed-7507245;			
RA	Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P., Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.,			
RT	"STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells."			
RT	"Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic cells."			
RL	Blood 82:1110-1119(1993).			
RN	[3]			
RP	SEQUENCE OF 783-942 FROM N.A.			
RC	TISSUE-Testis;			
RX	MEDLINE-91169547; PubMed-2004790;			
RA	Rosnet O., Mattei M.-G., Marchetto S., Birnbaum D.,			
RT	"Isolation and chromosomal localization of a novel FMS-like tyrosine kinase gene."			
RL	Genomics 9:380-385(1991).			
CC	-!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.			
CC	-!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- TISSUE SPECIFICITY: BONE MARROW CELLS.			
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.			
CC	-!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	-!- DATABASE: NAME=PROW; NOTE=CD guide CD135 entry;			
CC	WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd135.htm"			
CC	-!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;			
CC	WWW="http://www.infobioen.fr/services/chromancer/Genes/FLT3ID144.html"			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its			

34	827	15.7	821	1	FGR2_MOUSE	P21803 mus musculus
35	807	15.3	806	1	CEK2_CHICK	P18460 gallus gall
36	803.5	15.2	819	1	FGR1_CHICK	P21804 gallus gall
37	790.5	15.0	654	1	BFR2_HUMAN	Q01742 homo sapien
38	784	14.9	812	1	FGR1_XENLA	P22182 xenopus lae
39	779.5	14.8	822	1	FGR1_HUMAN	P11362 homo sapien
40	779	14.8	822	1	FGR1_MOUSE	P16092 mus musculus
41	766.5	14.5	806	1	FGR3_HUMAN	P22607 homo sapien
42	765	14.5	822	1	FGR1_RAT	Q04589 rattus norv
43	754	14.3	801	1	FGR3_MOUSE	O61851 mus musculus
44	728.5	13.8	1052	1	FGR2_DROME	Q09147 drosophila
45	716	13.6	1115	1	RET_MOUSE	P35546 mus musculus

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 002687; AAA18947.1; -
CC EMBL: 226652; CAA81393.1; -
CC EMBL: L36162; AAA35487.1; -
CC HSSP: P11362; 1FGK.
CC Genew: HGNC:3765; FLT3.
CC MIM: 136351; -
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003600; Ig_Like.
CC InterPro: IPR001824; RTK_kinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; Ig: 1.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 2.
CC SMART: SM00410; IG_Like; 1.
CC SMART: SM00219; Tyr_Kc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC Signal: Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
KW SIGNAL 1 26
FT CHAIN 27 993
FT DOMAIN 27 543
FT TRANSMEM 544 563
FT DOMAIN 564 993
FT DOMAIN 610 943
FT NP_BIND 616 624
FT BINDING 644 644
FT ACT_SITE 811 811
FT CARBOHYD 43 43
FT CARBOHYD 100 100
FT CARBOHYD 151 151
FT CARBOHYD 306 306
FT CARBOHYD 323 323
FT CARBOHYD 351 351
FT CARBOHYD 354 354
FT CARBOHYD 473 473
FT CARBOHYD 502 502
FT CARBOHYD 541 541
FT CONFLICT 8 8
FT CONFLICT 10 11
FT CONFLICT 78 78
FT CONFLICT 227 227
FT CONFLICT 346 346
FT CONFLICT 940 940
SQ SEQUENCE 993 AA; 112804 MW; 16790124B02F6BBF CRC64;

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVSAMIFGTITNQDLPIVKVCLINHKNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVSAMIFGTITNQDLPIVKVCLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRQSSGTVYAAAEVDVSAITLQVLVDAPGNISCLWFKHSSLNCCQPHFDL 120
Db 61 EDLGCALRQSSGTVYAAAEVDVSAITLQVLVDAPGNISCLWFKHSSLNCCQPHFDL 120

Qy 121 QNRGVVSMVILKMTETQAGYLLFIQSEATNYTLFTVSIIRNTLLYTLRPYFRKMNQD 180
Db 121 QNRGVVSMVILKMTETQAGYLLFIQSEATNYTLFTVSIIRNTLLYTLRPYFRKMNQD 180

Qy 181 ALVCISESPEPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISESPEPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

Qy 241 CTRLFTIDLNQTPOTTLPOLELKVCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
Db 241 CTRLFTIDLNQTPOTTLPOLELKVCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYE 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYE 360

Qy 361 IDQYEEFCFSVRFRKAYPOIRCTWTFSRKSFCPEQKGLDNGYSISKFCNKHKHPGEYI 420
Db 361 IDQYEEFCFSVRFRKAYPOIRCTWTFSRKSFCPEQKGLDNGYSISKFCNKHKHPGEYI 420

Qy 421 ENDDAQFTKMTLNIRRRKPOVLAESAASQSCFSGYPLPSWTWKCKSDKSPNCTEEITE 480
Db 421 ENDDAQFTKMTLNIRRRKPOVLAESAASQSCFSGYPLPSWTWKCKSDKSPNCTEEITE 480

Qy 481 GWNKRANKRVFGQWVSSSTLNSEAIKGLFKCCAYNSLGTSCETILLNSPGPPFIQD 540
Db 481 GWNKRANKRVFGQWVSSSTLNSEAIKGLFKCCAYNSLGTSCETILLNSPGPPFIQD 540

Qy 541 NISFYATIGVCLLFTIWLTLCHIKYKQFVYESOLQMVQVVTGSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFTIWLTLCHIKYKQFVYESOLQMVQVVTGSSDNEYFYVDFREYED 600

Qy 601 LKWEPPRENLEFGKVLGSCAFKVMNATAYGSKTCVSIQAVKMLKEKADSSERREALMS 660
Db 601 LKWEPPRENLEFGKVLGSCAFKVMNATAYGSKTCVSIQAVKMLKEKADSSERREALMS 660

Qy 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIFECYCCYDGLLNLRSKREKFRHTWTEIFKE 720
Db 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIFECYCCYDGLLNLRSKREKFRHTWTEIFKE 720

Qy 721 HNFSFYPTQSHPNMPSGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRLSEEDL 780
Db 721 HNFSFYPTQSHPNMPSGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRLSEEDL 780

Qy 781 NVLTFEDLLCFAYQVAKGMEFFLEKSCVHRDLAARNVLTGHKVVKICDFGLARDMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFFLEKSCVHRDLAARNVLTGHKVVKICDFGLARDMSDS 840

Qy 841 NYVVRNARLPVKWMAPESLFEGIYTIKSDVWSYIGLLMEIFSLGVNPPYGPVPDANFYK 900
Db 841 NYVVRNARLPVKWMAPESLFEGIYTIKSDVWSYIGLLMEIFSLGVNPPYGPVPDANFYK 900

Qy 901 LIQNGFKMDQPPFYATEEIIYIMOSWAFDSRKRSPFNLTSLGCOLADAEAEAMVQVNDG 960
Db 901 LIQNGFKMDQPPFYATEEIIYIMOSWAFDSRKRSPFNLTSLGCOLADAEAEAMVQVNDG 960

Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 2
FLT3_MOUSE STANDARD; PRT; 992 AA.
AC Q00342;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).
GN FLT3 OR FLT-3 OR FLK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91292518; PubMed=1648448;
RA Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;

RT "A receptor tyrosine kinase specific to hematopoietic stem and
RT progenitor cell-enriched populations.";
RL Cell 65:1143-1152(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92019834; PubMed=1656368;
RA Rosnet O., Marchetto S., Delapetiere O., Birnbaum D.,
RT "Mucine Flt3, a gene encoding a novel tyrosine kinase receptor of the
RT PDGFR/CSF1R family";
RL Oncogene 6:1641-1650(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93205405; PubMed=8384358;
RA Maroc N., Rottapel R., Rosnet O., Marchetto S., Lavezzi C.,
RA Mannoni P., Birnbaum D., Dubreuil P.;
RT "Biochemical characterization and analysis of the transforming
RT potential of the Flt3/Flk2 receptor tyrosine kinase.";
RL Oncogene 8:909-918(1993).
CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-
CC ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64689; AAA37634.1; -
CC EMBL; X59398; GAA42041.1; -
CC F1R; A39931; A39931.
CC HSP; P11362; IFGK.
CC MG; MGI:95559; Flt3.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003600; Ig_like.
CC InterPro: IPR001824; TykinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00410; Ig_like; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KTN_III; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
CC SIGNAL 1 27 POTENTIAL.
CC CHAIN 28 992 FL CYTOKINE RECEPTOR.
CC DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 545 564 POTENTIAL.
CC DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 611 946 PROTEIN KINASE.
CC NP_BIND 617 625 ATP (BY SIMILARITY).
CC BINDING 645 645 ATP (BY SIMILARITY).
CC ACT_SITE 814 814 BY SIMILARITY.
CC CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT

FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 150 150 R -> A (IN REF. 2).
FT CONFLICT 242 242 C -> S (IN REF. 2).
FT CONFLICT 726 726 S -> F (IN REF. 2).
FT CONFLICT 957 979 CIRTSLHLPKQAPQOORGGILRAO -> MYQNMGNVPEHPS
FT CONFLICT 983 983 IYQNRRLPSREAGSEPP (IN REF. 2).
FT CONFLICT R -> A (IN REF. 2).
SQ SEQUENCE 992 AA; 112639 MW; 407A08785372100 CRC64;

Query Match 84.0%; Score 4429.5; DB 1; Length 992;
Best Local Similarity 84.1%; Pred. No. 3.7e-289;
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MPALA-RDAGTVPLLVVSAMIFGTTITNODLPVTKVILNHKNNDSSVGSXPYPMVSES 59
DB 1 MRALQSRDRLLLVLSVMILETVTNQDLPVTKVILSHENNSSAGKPSYVMVRGS 60

QY 60 PEDLGCALRQSGTVEAAAEVDVSASITLQVLDVADPQNISCLWVFKHSSLNCQPHFD 119
DB 61 PEDLOCTPRQSGTVEAATVEAESGSTITLQVQLATPGDLSCLWVFKHSSLGCQPHFD 120

QY 120 LQNRGVSVVILKMTQAGEYLFFIQSEATNTYLTFTVSIKNTLLYLRPPYFKRMENQ 179
DB 121 LQNRGIVSMALLNVTVTQAGEYLLHQSEARNTYLTFTVNRDQTLVLRPPYFKRMENQ 180

QY 180 DALVCISESVPEPFWVLVDCSGESCKEESPAVAVKKEKVLHFLGTDIRCCARNELGR 239
DB 181 DALLCISEGPEPTVEWVLCSHRESCKEESPAVAVKKEKVLHFLGTDIRCCARNALGR 240

QY 240 ECTRLPTIDNPTQTLTQFLKVGEPWIRKAVVHNGHFGLTWELENKALEEGNYFE 299
DB 241 ECTKFTIDNQAOSTLQFLKVGEPWIRKAVHNGHFGLTWELENKALEEGSYFE 300

QY 300 MSTYSTNRTMIRLPAFVSVVARNDTGYTCSSSKHPSQSALVTIVKGFIGNATNSSEY 359
DB 301 MSTYSTNRTMIRLLAFVSVVGRNDTGYTCSSSKHPSQSALVTILEKGFIGNATNSSEY 360

QY 360 EIDYEEFCFSVRKAYPQIRCTWTFSRKSFCEQKGLDNGYSISFCNKHKHFGEYIFH 419
DB 361 EIDPYEKFCFSVRKAYPRICWTWFSQASFPCEQKGLDNGYSISFCNKHKHFGEYIF 420

QY 420 AENDDAQFTKMTLNTRRKPOVLAEASQASCFSDGYPLPSWTWKKCSKSPNCTEIT 479
DB 421 AENDDAQFTKMTLNTRRKPOVLAEASQASCFSDGYPLPSWTWKKCSKSPNCTEIP 480

QY 480 EGVWNRKANRKFVGOVSVSSTLNMSIAKGLVKCCAYNSLGTSCETILLNSPGPPFIQ 539
DB 481 EGVWNRKANRKFVGOVSVSSTLNMSIAKGLVKCCAYNSLGTSCETILLNSPGPPFIQ 540

QY 540 DNISFYATIGVCLLFTIVLTLTCHIKYKQFRYESOLQWQVVTGSSDNEYFYVDFREY 599
DB 541 DNISFYATIGVCLLFTIVLTLTCHIKYKQFRYESOLQWQVVTGSSDNEYFYVDFREY 600

QY 600 DLKWEFPRENLEFGKVLGSGAFKVMNATAYGISTKGTGVSQVAVKMLKEKADSSREALM 659
DB 601 DLKWEFPRENLEFGKVLGSGAFKVMNATAYGISTKGTGVSQVAVKMLKEKADSSREALM 660

QY 660 SELKMTQLGSHENIVNLLGACTLSGPYILFEYCYGDLNLYLRKREKFRHTWTIFEK 719
DB 661 SELKMTQLGSHENIVNLLGACTLSGPYILFEYCYGDLNLYLRKREKFRHTWTIFEK 720

QY 720 EHNFSYPTFQSHNSSMPGSRVQIHPDSDOIISGLHNGSHSEDETEYENQRL--EEE 777
DB 721 EHNFSYPTFQSHNSSMPGSRVQIHPDSDOIISGLHNGSHSEDETEYENQRLAEE 780

QY 778 EDNLVLTFFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLYTHGVKVKICDFGLARDIM 837
DB 781 EDNLVLTFFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLYTHGVKVKICDFGLARDIL 840

QY 838 SDSNVYVRGNARLPKVMAPESLFEIGIYTKSDVWSYIGILLMEIFSLGVNPNYPGIPVDAN 897


```

Db 841 SDSSVVRGNARLPVKNAPESLFGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVDAN 900
Qy 898 FYKLQNGFKMDQPPYATEEIIYIMQSWAFDSRKRPPNLTSLFGCOLAAEARMYQN 957
Db 901 FYKLQSGFKMEQPEYATEGIYVQSWAFDSRKRPPNLTSLFGCOLAAEAEAC--- 957
Qy 958 VDGVRVSECPHYQNRPPSRENDLGLLSPQAQVE 991
Db 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQRQVK 986

RESULT 3
KIT_CAPHI
ID KIT_CAPHI STANDARD; PRT; 978 AA.
AC Q28317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shiba; TISSUE=Cerebellum;
RX MEDLINE=97342548; PubMed=9199245;
RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,
RA Takamura Y., Sawasaki T., Tachi C.;
RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats
RT and a novel alanine insertion specific to goats and sheep in the
RT kinase insert region.";
RL Biochim. Biophys. Acta 1352:151-155(1997).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D45168; BAA08116.1; -
CC HSSP; P11362; 1FGK.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003600; Ig_Like.
CC InterPro; IPR001824; RTKinaseIII.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00410; IG_Like; 2.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00219; TyKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22
FT CHAIN 23 978
FT DOMAIN 23 521
FT TRANSMEM 522 544
FT DOMAIN 545 978
FT DOMAIN 590 939
FT NP_BIND 596 604
FT BINDING 624 624
FT ACT_SITE 794 794
FT MOD_RES 825 825
FT CARBOHYD 94 94
FT CARBOHYD 130 130
FT CARBOHYD 145 145
FT CARBOHYD 284 284
FT CARBOHYD 294 294
FT CARBOHYD 301 301
FT CARBOHYD 321 321
FT CARBOHYD 353 353
FT CARBOHYD 368 368
FT CARBOHYD 401 401
FT CARBOHYD 464 464
FT CARBOHYD 487 487
FT CARBOHYD 487 487
SQ SEQUENCE 978 AA; 109722 MW; CA4D663F98205CA9 CRC64;

Query Match 24.6%; Score 1295.5; DB 1; Length 978;
Best Local Similarity 33.3%; Pred. No. 3,9e-79;
Matches 335; Conservative 169; Mismatches 339; Indels 163; Gaps 35;

Qy 47 VGKSSYPVMSPELQGC-ALRPQSGTGYEAAVEVDVSASITLOLVDAQGNISCLW 105
Db 20 VQTGSSQPSV--SPGELSLPSIHPAKS-----ELIVSGDEIRLLCTDPGFVK--W 66
Qy 106 VFHSSLNCPHFDLQNRGVSMVLKMTQTQAGEYLLFIQSEAT---NYTLFTVSIRN 162
Db 67 TFE-----ILGOLSEKTPWEIT-EKAEATNTGNTCTNKGGLSS 105
Qy 163 TL-----LYTLRRPYFRKMENQDALVCISSEVPPIVEWVLCDSQGESKESPAV 213
Db 106 SIIVFVRDPEKFLIDLPLXGKEEN-DTLVRCPLTQPE-VTNYSLTGCEGKPLKDLTFV 163
Qy 214 -----VKKEKVLHLEFGTDIRC--CARNELGRE-CTRLFTID----LNQTPQTLP 258
Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANQKGSMLSKKETLKVRAAIKAVPVVSVS 217
Qy 259 Q--LFLKVGEPLEWIRKAVHVNHGFLTWELNKALEEGNYFEMSTYSTNTRMIRILFAF 316
Db 218 KTSYLLREGEEFAVTCILKDVSSVDSMMWIKENSQOSKAQTKNNSWHQGFYSYLRQERLT 277
Qy 317 VSSVARNDGYTCSSSKHPQSALVT---IVGKGFINA-----TNSSEDEYIDQY 364
Db 278 ISSARVNDSGVMFYANNTGSGANVTITLEVVDKGFNIPEPMNTTVFVNDGENVDL--- 334
Qy 365 EEFCSVRFKAYPQ-----IRCTWTFSRKFPCCQKGLDNGYSISKFCNKKH----- 411
Db 335 -----VVEEAYPKPEHRQWIYMNRTSTDKWDYPKSE-----NESNIRVYNELHLRLK 384
Qy 412 --QPGEYIFHAENDDAQFTKMTFLINIRRKPOVLAEASA--SQASCFSDGYPLPSWTKKC 467
Db 385 GTEGGTYTFHVSNDVSNSSVTFNVNNTKPEILTHDRLVNQMQLQCVAAAGPEPTIDWYFC 444
Qy 468 SDKSPNCTEEI-TEGVNRRKANRKFQGVWSSSTLNNSSEAIKGLFVKCCAYNSLGTSCET 526
Db 445 PGTEQRCSVPVGPVDVQIQNSVSPFGKLVYSTIDSTDKHNGTVECRAYNDVGKSSAS 504
Qy 527 ILLNSPGPFPIQDN-----ISFYATIGVCLLFIYVLTLLICHKYKKFY 572
Db 505 F-----NFAFKGNKEQIHAHTLFTPLLLIGFVIAAGLMCIFVMILT-----YKYLQPMY 554

```

QY 573 ESQOMV-QVTGSSDNEYFYVDFREYEDLKWFFPRENLEFGKVLGSGAFGKVMNATAYG 631
 Db 555 EYQWKVVEING---NNVYIDPTQPYDHKWFFPRNRLSFGKTLGAGAGKVVETAYG 611
 QY 632 ISKTGVSIOVAVKMLKERADSSERREALMSELKMMQTLGSHENIVNLLGACTLSGGPIYLIF 691
 Db 612 LKSDAATVAVKMLKPSAHLTREALSELKVLVSLGNHNMIVNLLGACTIGGPTLVIT 671
 QY 692 EYCCVGDLLNLYRSKREKFRHTWTE-----IFKEHNEFSYPTFQSHPNSSM---PGSRE 742
 Db 672 EYCCVGDLLNLYRRKDSFICKQEDHAVALYKLLHKSSESCNDSTWEYMDKPGVSY 731
 QY 743 VQIHPSDQISGLHGNHSEDEIEYENOKRLEEBEDLNLVTFEDLCLLFCAYQVAKGMEFL 802
 Db 732 VVPTKAADKRRSARIGSY-----IERDVTPTAIMEDEL-ALDLEDLLSFSYQVAKGMAL 785
 QY 803 EFKSCVHRDLAARNVLVTHGKVVKICDGLARDIMSDSNVYVGRNARLPVKWMAPESLFE 862
 Db 786 ASKNCIHRDLAARNLLTHGRITKICDGLARDINDSNVYVGNARLPVKWMAPESIEN 845
 QY 863 GYITIKSDWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGKMDOPFYATEEYIIM 922
 Db 846 CVYTESDWSYGIFLWELFSLGSSYPGMPVDSFYKMKIEGFRMLSPHAPAEYIM 905
 QY 923 QSCWAFDRKRSFPLNLSFGLGCOLADAEAMYNQYVGRVSEC-PH 967
 Db 906 KTCWDADLPKRTFTQVQIOLIEKQISESTNHIYSN---LANCSPH 947
 RESULT 4
 KIT_BOVIN
 ID KIT_BOVIN STANDARD; PRT; 977 AA.
 AC P43481.
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
 GN KIT.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94215924; PubMed=7512939;
 RA Kubota T., Hikono H., Sasaki E., Sakurai M.;
 RT "Sequence of a bovine c-kit proto-oncogene cDNA."
 RL Gene 141:305-306(1994).
 CC -!- GROWTH FACTOR. THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
 CC 3-KINASE (PI3K).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D16680; BAA04084.1;
 DR HSSP; P11362; IFKC.
 DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00410; Ig_Like; 2.
 DR SMART; SM00408; IgC2; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Proto-oncogene; tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 22
 FT CHAIN 23 977
 FT DOMAIN 23 521
 FT TRANSMEM 522 544
 FT DOMAIN 545 977
 FT DOMAIN 590 938
 FT NP_BIND 596 604
 FT BINDING 624 624
 FT ACT_SITE 793 793
 FT MOD_RES 824 824
 FT CARBOHYD 94 94
 FT CARBOHYD 130 130
 FT CARBOHYD 145 145
 FT CARBOHYD 284 284
 FT CARBOHYD 294 294
 FT CARBOHYD 301 301
 FT CARBOHYD 321 321
 FT CARBOHYD 353 353
 FT CARBOHYD 368 368
 FT CARBOHYD 401 401
 FT CARBOHYD 464 464
 FT CARBOHYD 487 487
 SQ SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;
 Query Match 24.4%; Score 1286; DB 1; Length 977;
 Best Local Similarity 33.2%; Pred. No. 1.7e-78;
 Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;
 QY 47 VQKSSYPMVSESPDLGC-ALRPQSSGYVEAAAVEVDVSASITLQVLVDAPGNISCLW 105
 Db 20 VQTGSSOPSV--SPGELSLSIHPAKS-----ELIVSGDEIRLLCTDPGFVK--W 66
 QY 106 VPKHSLNCQPHFDLONRGVSVMLKMTETQAGEYLLFTQSEAT---NYTILFTVSIRN 162
 Db 67 TFE-----ILGQISEKTNPEWIT-EKAEATNTGNYCTNKGGLSS 105
 QY 163 TL-----LYTLRRPVRKMNODALVCISEPPIVENVLDCDSQGESKEESPAY 213
 Db 106 SIYFVRDPEKFLDLDPLYGKEEN-DTLVRCPLTDPPE-VTNYSLTGCGEKLPLKDLTFV 163
 QY 214 -----VKKKEKVLHFGTDIRC--CARNELGRE-CTRFLTID---LNQTPQTLTP 258
 Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANORGKSMLSKFTLVKRAIKAVPVSVS 217
 QY 259 Q--LFLKVGEPILWIRKAVVHNVHFGLTWELENKALEEGNYFEMSTYSTNRTMIRILFAF 316
 Db 218 KTSYLLREGEFAVTCCLIKDVSSVSDSMWIKENSQOTKQTKKNSWHQGFVSLRQERLT 277
 QY 317 VSSVARNDGYTCSSSKIPSSQALVT---IVCKGGINA-----TNSSDEYIDQY 364
 Db 278 ISSARVNDGSGVFCYANNTFGSANVTTLLEVVDKGFINFPMMNTTVFVNDGENVDL--- 334
 QY 365 EEFCEFSVREKAVPO-IRCTWTFSRKS-----FPCQKGLDNGYSTSKFCNKHK----- 411
 Db 335 -----VVEYEPKPVHROWIYNNRTSTDKWDDYPKSE-----NESNIRYVNEHLHTRLK 384

```

Qy 412 --OPCEYTFHAENDDAQFTKMTFLNRKPOVLAESA--SQASCFSDCGYPLPSMTWKKC 467
Db 385 GTEGTYTFHVSNSDVNSVTFVNTKPEILTHRLVNGMLQCVAAAGFPPTIDWYFC 444
Qy 468 SDKSPNCTEEI--TEGVWNRKANRYGVQWSSSTLNMSPAIFGLFKCAYNSLGTSCFT 526
Db 445 PGTEQRCSPVGPVDPVQIONSSVPGKLVVYSTIDDSFKHNGTVCECRAYNDVCKSSAS 504
Qy 527 ILLNSPGFPP-----FIQDNISFYATIGVCLFIVVLITLLICHKKYKQFRESQLOM 578
Db 505 FNFAPKGNKSEQIHAHTLFTPLLLIGFVIAAGLMCFIVMLT----YKLOKPMYEVQWKV 560
Qy 579 V-OVTGSSDNEYFYDFEYEDLKWEPRENLEFGKVLGSAFGKVMNATAYGSKTGCV 637
Db 561 VEEING-----NNVYIDPTQLPYDHKWEPRNRLSFGKTLGACAFKGVVEATAYGLIKSDA 617
Qy 638 SIQVAVKMLKERKADSSEREAALMSLKMOTLGSHEINVLGACTLSGPIYLIFFEYCCYG 697
Db 618 AMTVAVKMLKPSAHLTEREALMSLKVLSYLGHNHNVNLLGACTTIGGPTLVITEYCCYG 677
Qy 698 DLLNVLRSKRKFHTWTE-----IFKEHNSFTPTFOSHNSM---PGSREVOIHDP 748
Db 678 DLLNVLRSKRKFHTWTE-----IFKEHNSFTPTFOSHNSM---PGSREVOIHDP 748
Qy 749 SDOISLHGNSFSEDEIEYENQKLEEEEDNLVTFEDLLCFAYQVAKGMFLEPKSCV 808
Db 737 ADKRSARIGST-----IERDTPAIMEDEL-ALDELLSFSQVAKGMFLEPKSCV 808
Qy 809 HRDLAARNVLVTHGVKVKICDFGLARDIMSDSNVYVGNARLPVKNWMAFESLFEGIYTIK 868
Db 791 HRDLAARNVLVTHGVKVKICDFGLARDIMSDSNVYVGNARLPVKNWMAFESLFEGIYTIK 868
Qy 869 SDVWSYGILLWEIFSLGVNPNYGPVDPANFYKLQNGFKMDQFPYATEEIIYIMQSCWAF 928
Db 851 SDVWSYGIFLWELFSLGSSPYPCMPVDSFYKMKIEGFRMLSPHAPAEWYDITKTCWDA 910
Qy 929 DSRKRPSPNLTSLGCOLADAEAEAMVQVNDGRVSEC-PH 967
Db 911 DPLKRPTRKQIVOLIEKQISESTNHHYSN-----LANCSPH 946

RESULT 5
KIT_MOUSE
ID KIT_MOUSE STANDARD: PRT: 975 AA.
AC P05332; Q61415; Q61416; Q61417;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
GN KIT OR SL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=BALE/c;
RX MEDLINE=88296403; PubMed=2456920;
RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddle F.H.,
RA Besmer P.;
RT "Primary structure of c-kit: relationship with the CSF-1/PDGF
RT receptor kinase family -- oncogenic activation of v-kit involves
RT deletion of extracellular domain and C terminus.";
RL EMBO J. 7:1003-1011(1988).
RN [2]
RP SEQUENCE FROM N.A. (TRUNCATED FORM).
RC STRAIN=ICR;
RX MEDLINE=92331813; PubMed=1378413;
RA Rossi P., Marziani G., Albanesi C., Charlesworth A., Geremia R.,
RA Sorrentino V.;
RT "A novel c-kit transcript, potentially encoding a truncated receptor,
```

```

RT originates within a kit gene intron in mouse spermatids.";
RL Dev. Biol. 152:203-207(1992).
RN [3]
RP LIGAND.
RX MEDLINE=91006023; PubMed=1698611;
RA Tan J.C., Buck J., Levi E., Besmer P.;
RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a
RT fibroblast derived growth factor stimulates mast cells and erythroid
RT progenitors.";
RL EMBO J. 9:3287-3294(1990).
RN [4]
RP VARIANT W42 ASN-790.
RX MEDLINE=90100577; PubMed=1688471;
RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;
RT "The dominant W42 spotting phenotype results from a missense mutation
RT in the c-kit receptor kinase.";
RL Science 247:209-212(1990).
RN [5]
RP VARIANTS W37 LYS-582; W41 MET-660 AND W41 MET-831.
RX MEDLINE=90269214; PubMed=1693331;
RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,
RA Besmer P.;
RT "Molecular bases of dominant negative and loss of function mutations
RT at the murine c-kit/white spotting locus: W37, W41 and W42";
RL EMBO J. 9:1805-1813(1990).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DISEASE: WHITE-SPOTTING VARIANT INDUCES SEVERE EFFECTS ON
CC PIGMENTATION, GAMETOGENESIS AND HEMATOPOIESIS. MICE HOMOLOGOUS
CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00864; CAA68772.1; -
CC EMBL; X65997; CAA46798.1; -
CC EMBL; X65998; CAA46799.1; ALT_SEQ.
CC EMBL; X65998; CAA46800.1; -
CC PIR; S00474; TVMSKT.
CC HSSP; P11362; IFGK.
CC MGD; MGI:96677; Kit.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003600; Ig_Like.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00410; IG_Like; 3.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW
```

Immunoglobulin domain; Disease mutation.

1 22
KW SIGNAL 23 975 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
FT CHAIN 23 519 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 520 542 POTENTIAL.
FT TRANSMEM 543 975 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 588 935 PROTEIN KINASE.
FT NP_BIND 594 602 ATP (BY SIMILARITY).
FT BINDING 622 622 ATP (BY SIMILARITY).
FT ACT_SITE 790 790 BY SIMILARITY.
FT MOD_RES 821 821 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 1 12 MRGARGAWDLIC -> MAVAVFPFLPQQ (IN TRUNCATED FORM).
FT VARIANT 13 785 MISSING (IN TRUNCATED FORM).
FT VARIANT 582 582 E -> K (IN W37 SPOTTING).
FT VARIANT 660 660 T -> M (IN W3 SPOTTING).
FT VARIANT 790 790 D -> N (IN W42 SPOTTING).
FT VARIANT 831 831 V -> M (IN W41 SPOTTING).
SQ SEQUENCE 975 AA; 109001 MW; BAB5CA4D9AF9CD2A CRC64;

Query Match 24.0%; Score 1266; DB 1; Length 975;
Best Local Similarity 33.2%; Pred. No. 3.7e-77;
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

QY 35 CVLLNHKNDSSVGVSSSPVSVSPEDLGCALRQSSGTVYEAARVVDVVSASITLQVL 94
DQ 12 CVLLVLLRGQTSPSPASPGSPSP-----SIHPAQSS-----ELIVEAGDTL---- 54
QY 95 VDAPNISCL-----WVFKHSSLNCQPHFDLQNRGVSMVILKMTQAGYLLFIQSE 148
DQ 55 -----SLTCIDPDFVFWTFK-----YFVN-----ENVENKKNWQI-ERAE 89
QY 149 ATNVTILTSTIRNL-----LYTLRRPYFRKMNQDALVCISSPVPEIVE 195
DQ 90 ATR-TGTYTCSNGLTSSIIYVFRDPAKFLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146
QY 196 WVLCDSSQSCKEESPAV-----VKKEKVLHELFGTDIRCCARNELGRETRL--- 244
DQ 147 YSLIECDGKSLPTDLTFVPNPKAGITIKVRAYRLC---VRCQAQ-----RDGTWLHSD 199
QY 245 -FTIDLNOT---POTTLPOL--FLKVGELWIRCKAVHNVHGFLTW-----E 286
DQ 200 KFTLVRAIKAIPIVSVVPETSHLLKKGDTFVVCVTIKDVSTSVNSMWLKMNPQOHIQA 259
QY 287 LENKALEBGNFYEMSTYNTMTIRILFAFVSSVARNDTGYTSCSSKHPSSQALVT--- 343
DQ 260 VKHNSWHRGDF-----NYERQETLT-----ISSARVDSGVFPMCYANNTFGSANVTTLK 309
QY 344 IVGKGFINATN-SSDYEDIDQVEECFVSFRKAYPQ-TRCTWTFPSRKSPCQKGLD--- 398
DQ 310 VVEKGFINISPVKNTTTFVTVDGENVDLVVEYEAYPEKPHQOQWYMNRT--SANKGKDYVK 367
QY 399 -NGYSISKFCNKH-----OPGEYIFHAENDDAQTKMFTLNIRRRKPOVLA--EASAS 448
DQ 368 SDNKNIRYVQLRULTRUKGTEGGTYTLVNSDSASAVTFNYYNTPEILTYDRLING 427
QY 449 QASCFSDGYPSPWTKKCSKSPNCTBEITE-GVWNRKANKRVFGQWVSSSTLNMSAI 507
DQ 428 MLCQVAEGFPETIDWYECTGAEOQCTTPVSPDVQVQNVSVSPFGKLVVQSSIDSSVFR 487
QY 508 KGLVKKCAYNLSLGTSCETILLNSPGPPFF-----IQDN-----ISFYATIGVCLLFIW 557
DQ 488 HNGTVECKASNDVGRKS--SAFFN-----FAFREQIQIAHLFTPLLLIGFVVAAGAMGIIVW 541
QY 558 LTLILCHYKQKFRYESOLQMV-QVTGSSDNEYFYVDFREYEDYDLKWEFPRENLEFGKVL 616

Db 542 LT-----YKYLQKPMYEVQWKVVEEING---NNYVYIDPTQLPYDHKWEFPRNLSFGKTL 594
QY 617 GGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREAALMSKMMTQOLGSHENIVN 676
DQ 595 GAGAFKVVYATAYGLIKSDAAMTAVKMLKPSAHLTREALMSKLVLSYLGNNHMINV 654
QY 677 LLGACTLSGPVILYIFCYCCYDGLLNLVLRKREKF-----HRTWTEIFKEHNEFSYPTFQ 730
DQ 655 LLGACTVGGTTLVITEYCYCYDGLLNLVLRKRDSEIFSKQEEQAAALYKLLHSTPESCD 714
QY 731 SHPNSSM---PGSREVQIHPDSQISGLHGNSPHSEDEIEYENOKLEEDENLVTFED 787
DQ 715 S-SNEYMDMKPGVSIV-VPTKTDKRSARDSY-----IERDVTPALMEDDEL-ALDLD 766
QY 788 LLCFAYQVAKGMEFLFECVHRDLAARNVLYTHGVKVKICDFGLARDINSDSNVYVRGN 847
DQ 767 LLSFSYQVAKAMAFASKNCIHRDLAARNILLTHGRITTKICDFGLARDINSDSNVYVRGN 826
QY 848 ARLPVKWMAPESEFEGIYTIKSDVWSYGGILLWEIFSLGVNYPGIPVDANFYKLIQNGFK 907
DQ 827 ARLPVKWMAPESEFSCVYTFESDVWSYGFILWELFSLGSSPYGMPVDSFYKMKREGFR 886
QY 908 MDQPFYATEEIIYIMOSWAFDSRKRSPNLTSLFCOLADAEAMYQNY 958
DQ 887 MYSPEHAPAEYDMVKTCWDADPLKRPPTKQVVLIEKQISDSTKHIYSNL 937

RESULT 6
KIT_HUMAN
ID KIT_HUMAN STANDARD; PRT; 976 AA.
AC P10721.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine kinase protein kinase kit) (c-kit) (CD117 antigen).
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=term placenta, and fetal brain;
RX MEDLINE=88111521; PubMed=2448137;
RA Yarden Y., Kung W.-J., Yang-Feng T., Coussens L., Munemitsu S.,
RA Duli T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;
RT "Human proto-oncogene c-kit; a new cell surface receptor tyrosine
RT kinase for an unidentified ligand.";
RL EMBO J. 6:3341-3351(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93064697; PubMed=1279499;
RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;
RT "Organization and nucleotide sequence of the human KIT (mast/stem
RT cell growth factor receptor) proto-oncogene.";
RL Oncogene 7:2207-2217(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
RN [4]
RP VARIANT LYS-583.
RX MEDLINE=92291284; PubMed=1376329;
RA Fleischman R.A.;
RT "Human piebald trait resulting from a dominant negative mutant allele
RT of the c-kit membrane receptor gene.";
RL J. Clin. Invest. 89:1713-1717(1992).

[5]
 RP VARIANT LEU-584.
 RX MEDLINE=92133600; PubMed=1370874;
 RA Spritz R.A., Giebel L.B., Holmes S.A.;
 RT "Dominant negative and loss of function mutations of the c-kit
 RT (mast/stem cell growth factor receptor) proto-oncogene in human
 RT piebaldism.";
 RL Am J. Hum. Genet. 50:261-269(1992).
 [6]
 RN VARIANT ARG-664.
 RP MEDLINE=92020918; PubMed=1717985;
 RA Giebel L.B., Spritz R.A.;
 RT "Mutation of the KIT (mast/stem cell growth factor receptor)
 RT proto-oncogene in human piebaldism.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
 [7]
 RN VARIANT VAL-816.
 RP MEDLINE=94013473; PubMed=7691885;
 RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,
 RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,
 RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;
 RT "Identification of mutations in the coding sequence of the proto-
 RT oncogene c-kit in a human mast cell leukemia cell line causing
 RT ligand-independent activation of c-kit product.";
 RL J. Clin. Invest. 92:1736-1744(1993).
 [8]
 RN VARIANT'S PIEBALDISM GLY-791 AND VAL-812.
 RP MEDLINE=93322624; PubMed=7687267;
 RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;
 RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)
 RT proto-oncogene in human piebaldism.";
 RL J. Invest. Dermatol. 101:22-25(1993).
 [9]
 RN VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.
 RP MEDLINE=96287384; PubMed=8680409;
 RA Riva P., Milani N., Gandolfi P., Larizza L.;
 RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large
 RT Italian kindred with piebaldism.";
 RL Hum. Mutat. 6:343-345(1995).
 [10]
 RN VARIANT GIST VAL-559 DEL.
 RP MEDLINE=98361155; PubMed=9697690;
 RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,
 RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,
 RA Kitamura Y.;
 RT "Familial gastrointestinal stromal tumours with germline mutation of
 RT the KIT gene.";
 RL Nat. Genet. 19:323-324(1998).
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
 CC 3-KINASE (PI3K).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL
 CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC
 CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF
 CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.
 CC -!- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL
 CC STROMAL TUMOR (GIST).
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide Cd117 entry;
 CC WWP="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm"
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. haematol.;
 CC WWP="http://www.infobiogen.fr/services/chronocancer/Genes/KITID127.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X06182; CAA29548.1; -
 DR EMBL; X69301; CAA49159.1; -
 DR EMBL; X69302; CAA49159.1; JOINED.
 DR EMBL; X69303; CAA49159.1; JOINED.
 DR EMBL; X69304; CAA49159.1; JOINED.
 DR EMBL; X69305; CAA49159.1; JOINED.
 DR EMBL; X69306; CAA49159.1; JOINED.
 DR EMBL; X69307; CAA49159.1; JOINED.
 DR EMBL; X69308; CAA49159.1; JOINED.
 DR EMBL; X69309; CAA49159.1; JOINED.
 DR EMBL; X69310; CAA49159.1; JOINED.
 DR EMBL; X69311; CAA49159.1; JOINED.
 DR EMBL; X69312; CAA49159.1; JOINED.
 DR EMBL; X69313; CAA49159.1; JOINED.
 DR EMBL; X69314; CAA49159.1; JOINED.
 DR EMBL; X69315; CAA49159.1; JOINED.
 DR EMBL; X69316; CAA49159.1; JOINED.
 DR EMBL; U63834; AAC50968.1; -
 DR PIR; S01426; TVHUKT.
 DR HSSP; P11362; IFGK.
 DR Genew; HGNC:6342; KIT.
 DR MIM; 164920; -
 DR MIM; 172800; -
 DR MIM; 606764; -
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003006; Iq_MHC.
 DR InterPro; IPR003598; Iq_c2.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; TyrKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; Disease mutation.
 KW Immunoglobulin domain; Disease mutation.
 FT SIGNAL 1 22
 FT CHAIN 23 976 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
 FT DOMAIN 23 520 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 521 543 POTENTIAL.
 FT DOMAIN 544 976 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 589 937 PROTEIN KINASE.
 FT NP_BIND 595 603 ATP (BY SIMILARITY).
 FT BINDING 623 623 ATP (BY SIMILARITY).
 FT ACT_SITE 792 792 BY SIMILARITY.
 FT MOD_RES 823 823 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 283 283 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 320 320 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 559 559 MISSING (IN GIST).
 FT VARIANT 583 583 /FTid=VAR_007965.
 FT VARIANT 584 584 E -> K (IN PIEBALDISM).
 FT VARIANT 584 584 F -> L (IN PIEBALDISM).
 FT VARIANT 664 664 /FTid=VAR_004105.
 FT VARIANT 664 664 G -> R (IN PIEBALDISM).

FT	FT	791	791	/FTid=VAR_004106..	
FT	FT			R -> G (IN PIEBALDISM).	
FT	FT	812	812	/FTid=VAR_004107	
FT	FT			G -> V (IN PIEBALDISM).	
FT	FT	816	816	/FTid=VAR_004108..	
FT	FT			D -> V (IN MAST CELL LEUKEMIA;	
FT	FT			CONSTITUTIVELY ACTIVATED).	
FT	FT	893	896	/FTid=VAR_004109.	
FT	FT			MISSING (IN SEVERE PIEBALDISM).	
FT	FT	976	976	/FTid=VAR_004110.	
FT	FT			SEQUENCE 976 AA; 109864 MW; 8180CD/6817F3454 CRC64;	
Query Match					23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity					32.7%; Pred. No. 3.8e-76;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps					34;
QY	47	VKSSYPWSPEDLGCALRQSSGTVYEAAREVDSASTLQVLVDAPGNISCLWV	106		
DB	20	VQTGSSQPSVPG-EPSPPSIHFGKSDLI-----VRVGEIRLLCTDTPGFVK-WT	67		
QY	107	FKHSSLCNCPHFDLONRGVSVMLKMTETOAGEYLLFIOSEATN---YTLFTVSTINT	163		
DB	68	FE-----ILDETENKNQEWIT-EKAEATNTGKYCTTNKHGLSNS	106		
QY	164	L-----LYTLRRPYFRKMENQDALVCISBPPEIWEVWLCDQSGESCKEE-----	209		
DB	107	IYFVRDPAKFLVDRSLYCK-EDNTLVRCPLTDPE-VTNYSLKCGQKPLPKDLRFIP	164		
QY	210	----SPAVKKEKVLHELFGTDIRCCARNELGRECTRLFTIDL-----NOTPQTTLPO--L	260		
DB	165	DPKAGIMIKSVKRAYHRLC---LHCSVDQCKSVLSEKFLTKVRPAKPAVPVSVSKASY	221		
QY	261	FLKVGPELWIRKAVHVNHCFLTWELN---KALBEGVNFEMSTVSTNRTMIRILFAFV	317		
DB	222	LUREGEEFTVCTIKOVSSVSTWRENSQTKLOEKYNSWHHGDGFNYEQAT-----LTI	277		
QY	318	SSVARNDTGYTCCSSKHPSQSALVT---IVGKGFNA-----TNSSEYEDIOYE	365		
DB	278	SSARVNDSGVFMCYANTGCSANVTITLWVDKGFNFIPMINTTVFVNDGENVDL----	333		
QY	366	EPCFSVRKAYPO-ICTWTFTSRKF--PCEQKGLONGYSISKFNHK-----OPGE	415		
DB	334	----IVEYEAFPKPEHQWIYMNRTDKWEDYPKSESNIRYVSELHLTRLKGTGGT	389		
QY	416	YIFHAENDAOFTKMTLNIARRKPOVLA--EASASQASCFSDGYPPLPSWTWKCSKSPN	473		
DB	390	YTLVSNSDVNAAIAPNFVYNTKPEILTVDRLVNGMLQCVAAGPEPTIDWYFCPGTEQR	449		
QY	474	CTEETPE-GVWNRKANRKYFGOWSVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSP	532		
DB	450	CSASVLPVDVQTLNNSGPPFGKLVQSSIDSSAFKNGIVECKAYNDVGKT--SAYFN--	505		
QY	533	GPFPFTQDN-----ISFYATIGVCLLFIVVLTLTLLCHIKYKKOFRYESOLOM	578		
DB	506	--FAFGKNNKEOIHPHTLFTPLLIGFVIVAGMCIIVMLIT-----YKLOKPMYEQWKV	559		
QY	579	V-QVTGSSDNEFYVDFREYEDLKWEPFRENLEFGKVLGSGAFGVMNATYGISKTGV	637		
DB	560	VVEING---NNTVYIDPTQLPDYDHWKFEPRNRLSFKTLGAGAFGRVWEATAYGLKSDA	616		
QY	638	STQVAVKMLKEKADSSERELMSKMMQTLGSHENIVNLGACTLSGPIYLIFYECCYCG	697		
DB	617	AMTVAVKMLKPSAHLTERALMSKLVLSYLGNNHMINVLLGACTTGGPFLVITYECCYCG	676		
QY	698	DLNLVLRKREKPHRTWTE-----IFKEHNFSEFYPTQFQSHPNSSM---PGSREVOIHPD	748		
DB	677	DLNLVLRKRODSFICKQEDHAEALYKLLHKSCKSSDSTINEYMDKPGVSIV-VPTK	735		
QY	749	SDQISGLHGNSHSDEIEYENOKRLEEDENLVLTFFDLCLFAYQVAKMGMEFLFKSCV	808		
DB	736	ADKRRSVRIGSY-----IERDVTPTALMEDDEL-ALDLELLSFSYQVAKGMFLASKNCI	789		
QY	809	HRDLAARNVLVTHGKVVKTCDFGLARDINSDSNVYVGRNARLPVKWMAPESLFEGYITIK	868		

Db	790	HRDLAARNILLTHGRITKICDFGLARDIKNDSDNYVVKGNARLPVKWMAPESEIFNCVTFE	849		
QY	869	SDVWSYGILLWEIFSLGPNVPGIPVDANFYKLTQNGKMDQPFYATEEIIYIWSQWAF	928		
DB	850	SDVWSYGIFLWEFLSLGSSPYGMPVDSKFKYMKIKEGFRMLSPHAPAEYMDIMTKWDA	909		
QY	929	DSRRKPSFPNLTSLGCOLADAEEAMYNV	958		
DB	910	DPLARPTFKQIVQLIEKQISESTNHIYSNL	939		
RESULT 7					
KIT_CANFA					
ID	KIT_CANFA	STANDARD;	PRT;	975	AA.
AC	Q97799;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCPR)				
DE	(Proto-oncogene tyrosine-protein kinase Kit) (c-kit).				
GN	KIT.				
OS	Canis familiaris (Dog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
OX	NCBI_TaxID=9615;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99142897; PubMed=9989791;				
RA	Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;				
RT	"Clustering of activating mutations in c-KIT's juxtamembrane coding				
RT	region in canine mast cell neoplasms.";				
RL	J. Invest. Dermatol. 112:165-170(1999).				
CC	-I- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL				
CC	GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.				
CC	BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND				
CC	ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL				
CC	3-KINASE (PI3K).				
CC	-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein				
CC	tyrosine phosphate.				
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY-OF TYROSINE-				
CC	PROTEIN KINASES.				
CC	-I- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.				
This SWISS-PROT entry is copyright. It is produced through a collaboration					
between the Swiss Institute of Bioinformatics and the EMBL outstation -					
the European Bioinformatics Institute. There are no restrictions on its					
use by non-profit institutions as long as its content is in no way					
modified and this statement is not removed. Usage by and for commercial					
entities requires a license agreement (See http://www.isb-sib.ch/					
or send an email to license@isb-sib.ch).					
CC	EMBL; AF044249; AAD02327.1;				
DR	HSSP; P11362; IFGK.				
DR	InterPro; IPR000719; Euk_pkinase.				
DR	InterPro; IPR003599; Ig.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR003600; Ig_like.				
DR	InterPro; IPR001824; RTKinaseIII.				
DR	InterPro; IPR001245; Tyr_pkinase.				
DR	Pfam; PF00047; ig; 2.				
DR	Pfam; PF00069; pkinase; 1.				
DR	ProDom; PD000001; Euk_pkinase; 2.				
DR	SMART; SM00409; IG; 2.				
DR	SMART; SM00410; IG_like; 1.				
DR	SMART; SM00219; TyrKc; 1.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.				
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.				
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;				
KW	Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;				

[illegible]

DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00410; IG_like; 4.
 DR SMART; SM00408; IG2; 1.
 DR SMART; SM00219; TykKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 19
 FT CHAIN 20 980
 FT DOMAIN 20 509
 FT TRANSMEM 510 535
 FT DOMAIN 536 980
 FT DOMAIN 24 104
 FT DOMAIN 107 197
 FT DOMAIN 204 298
 FT DOMAIN 299 397
 FT DOMAIN 398 502
 FT DOMAIN 579 908
 FT NP_BIND 585 593
 FT BINDING 613 613
 FT ACT_SITE 776 776
 FT DISULFID 42 84
 FT DISULFID 127 177
 FT DISULFID 224 278
 FT DISULFID 417 482
 FT MOD_RES 807 807
 FT CARBOHYD 45 45
 FT CARBOHYD 73 73
 FT CARBOHYD 94 94
 FT CARBOHYD 153 153
 FT CARBOHYD 275 275
 FT CARBOHYD 286 286
 FT CARBOHYD 302 302
 FT CARBOHYD 335 335
 FT CARBOHYD 410 410
 FT CARBOHYD 477 477
 FT CARBOHYD 490 490
 SQ SEQUENCE 980 AA; 108506 MW; 4E5CF661E97CFCFF CRC64;
 Query Match 23.2%; Score 1225.5; DB 1; Length 980;
 Best Local Similarity 33.5%; Pred. No. 1.9e-74;
 Matches 321; Conservative 168; Mismatches 321; Indels 149; Gaps 34;
 QY 83 VDVSASITLQVLVDAPGNISCLW---VFKHSSILNCP-----HFDLQNRGVVSMVIL 131
 DB 32 VEPGTVTLRCV-----NGSVEDWDGPISPHWNLDLDPSSILTNATFQNTGYHCT-- 85
 QY 132 KMTETQAGEYL--LFIQSEATNYTL---FTVSIRNTLLTLRRPYFRKMNQDALV-CI 185
 DB 86 EPGNGGQGNATHLVKDPAPKPKVLAQEVTV-----LEGQDALLPCL 128
 QY 186 SESVPEPIVE--WVLCDGSGESCKEE-----SP---AVVKKEKVLHFLFGDIRCCARN 235
 DB 129 ---LTDPALEAGVSLVRGRVRLPQNTYNSFSPWHGFTTHKAKFIENHYV---QCSARV 181
 QY 236 ELGRECTRL-----FTIDLNOTPQTL-LPQLFKV-GEPLWIRCKAVHNVHGFGLTWLE 288
 DB 182 D-GRVTVSMGIWKLKQKDLGSGPATITLEPAELVIRQGEAAQVCSANIDVNFV----- 235
 QY 289 NKALEGN-----YFEMSTYSTNTRMIRILFAFVSSVARNDTGYTCSSSK---HPSQSAL 341
 DB 236 ---SLRHGDTKLTISQSDFDHNRQY-KVLTNLHDVSVFQDAGNYSCTATNANGNHSASV 292
 QY 342 VTIVGKGGINATN--SSEYDEIDQEEFCFSVRFKAYPQIRC-TWTFSRKSFCEQKGLD- 398
 DB 293 FRVSEAVYLNLTSEQSLLQEVTVGKVDLQVKVEAYPGLESEFNWTV-LGPFSDYQDKLDF 351
 QY 399 ----NGYSISKFCN----KKHQPGEYIFHAENDDAQFTKMTFLNTRKPKQVLAESASQA 450

DB 352 VTIKDTYRYVTSTLSLPRLKRSEAGRYSFLLARNAGGONALFELTLRYPPEVRVTMTLNG 411
 QY 451 S-----CFSDGYPLPSTWTKKCDKSPNCTEE---ITEGVNKRANRANRKFQWVSSSTLNM 503
 DB 412 SDTLLCEASGYPOPSVTWQCRSHTRDCDESAGLVLEDSEVLSQVPPHEVIVHSLAI 471
 QY 504 SEAIKGLFKCCAYNSIGSCETILLNSPGPPFFFDIODNISFYATIGVC---LLFIVVLT 560
 DB 472 GTLEHRTTECFRAFNSVGNSSOTFWPISIAHTQLPDELLFTPVLLTCMSIMALLLLLL 531
 QY 561 LICHYKKGKFRYESQLQMVQVTGSSDNEYFYVDFREYEDLKWEFFRENLEFSGKVLGSGA 620
 DB 532 LLLYKYQKPKYQVRWKIIE--SYEGNSYTFIDPTQLPYNEKWEFFPNNLQFGKTLGAGA 589
 QY 621 FGKVMNATYIGSKTGVSTQVAVKMLKEKADSSERELMSELKMMTOLGSHENTVNLGA 680
 DB 590 FGKVEATAFGLKEDAVLKVAVKMLKSTAHADKEKALMSELKIMSHLGQHENIVNLGA 649
 QY 681 CTLSGPIYLIFFEYCCYCDLLNLYRSKRE-----KFRWTWTEIFKE 720
 DB 650 CTGGGPVLVITEYCCYCDLLNLRQAAMLPGLSLSVGQDPEAGAGYKNHLEKKYVRD 709
 QY 721 HNFSFYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDEIEYENOKRLEEEDL 780
 DB 710 SDFS-----SQVDYTVEMRPVSTSSNDSFSEEDLCKEDGRPLE----- 749
 QY 781 NVLTFEDLLCFAYQVAKMEFFLEKSCVHRDLAARNVLTHTGKVKVTCDFGLARDIMSDS 840
 DB 750 ----LRDLLHFSQVAGQMAFLASKNCIHRDVAARNVLTSGRVAKIGDFGLARDIMSDS 805
 QY 841 NYVVRGNARLPVKWMAPESLFEGIYIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
 DB 806 NYIVKGNARLPVKWMAPESLFDCVYTVQSDVWSYGILLWEIFSLGLNPNYPGILNRSFYK 865
 QY 901 LIQNGFKMQDPFYATIEIYIMOSWAFDSKRKPSPLNTSLGCOL-ADAEAMYNQV 958
 DB 866 LVKDGQMAQAPAPAPAKNIYSIMQACWALEPTRRPTQQICSLQKQAEQDRRVPNYTNL 924
 RESULT 9
 ID KIT_FELCA STANDARD; PRT; 978 AA.
 AC Q28889;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
 GN KIT.
 OS Felis silvestris catus (cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RX MEDLINE=95140426; PubMed=7530827;
 RA Herbst R., Munemitsu S., Ullrich A.;
 RT "Oncogenic activation of v-kit involves deletion of a putative
 RT tyrosine-substrate interaction site.";
 RL Oncogene 10:369-379(1995).
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
 CC 3-KINASE (PI3K).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S76596; AAB33207.1; -
 HSSP; P11362; IFKC.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR001824; TyKinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00409; Ig; 2.
 DR SMART; SM00410; Ig_Like; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 22
 FT CHAIN 23 978
 FT DOMAIN 23 521
 FT TRANSMEM 522 544
 FT DOMAIN 545 978
 FT DOMAIN 590 938
 FT NP_BIND 596 604
 FT BINDING 624 624
 FT ACT_SITE 793 793
 FT MOD_RES 824 824
 FT CARBOHYD 94 94
 FT CARBOHYD 145 145
 FT CARBOHYD 284 284
 FT CARBOHYD 294 294
 FT CARBOHYD 301 301
 FT CARBOHYD 321 321
 FT CARBOHYD 353 353
 FT CARBOHYD 368 368
 FT CARBOHYD 401 401
 FT CARBOHYD 464 464
 FT CARBOHYD 487 487
 SQ SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;
 Query Match 23.2%; Score 1224.5; DB 1; Length 978;
 Best Local Similarity 32.9%; Pred No. 2.3e-74;
 Matches 325; Conservative 167; Mismatches 352; Indels 145; Gaps 35;
 QY 47 VGGSSYPVMSPELDC-ALRPQSSGTYYEAAAVEVDVSASITLQVLVDAPGNISCLW 105
 DB 20 VQTGSSQP--SASPGWSLPSIHPATS-----ELIVSAGDEIRLLCTDGFVK--W 66
 QY 106 VFKHSLNCQPHFDLQNRGWSVILKWTETQAGEYLLFIQSEAT---NYTIL----- 155
 DB 67 TFE-----TLQSSSEITHNEWIT-EKAEATNTGNYTCTNGGGLSS 105
 QY 156 -FTVSTRN-TLLYTLRRPYRKMNODALVCISESPPIVEWVLDCSOGESCKEESPAV 213
 DB 106 STYFVRDPAKFLVDLPYLGK-EDHDTLVRCPDTPDE-VTNISLRCGCEKPLPKDFTV 163
 QY 214 -----VKKEKVLHFGTDIRCA-----RNLGRECTRLFTIDLNPQTTLPLQL- 260
 DB 164 TDPKAGITIRNVKREYHRLC---LHCSADRKGSVLSKFKTLKVRRAIRAVPVVSYSKAS 220

QY 261 -FLKVGEPWIRCKAVHVNHGFLTWELNKALEGNYFEMSTYSTRNTRMIRILFAVSS 319
 DB 221 HLLREGEESVCLIKDVSSVDSSMIKENSQTNAPQSNSHHQGDFNFVROERUTISS 280
 QY 320 VARNDTGYTCSSSHPSOSALVT---IVKGGINA-----TNSSEYDIDQYEEF 367
 DB 281 ARVNDGSGFMCYANNTFFGSANVTTLLEVAKGIFPMMNTTIFVNDGENVDL----- 334
 QY 368 CFSVRPKAYPQ-----IRCTWTFSRKSPFCEQKGLDNGYSISKFCNKKH-----Q 412
 DB 335 --IYEYEAYPEKHORWVYNNRRLTKWEDYPKS-----DNESNI-RYVSELHTRLKGN 387
 QY 413 PGEYIFHAENDDAQFTKMETLNIRRRPQVLAES--ASQASCFSDGYPLPSWTWKKCSDK 470
 DB 388 GGTTFQVSNDSVNSVTLNYYVNTKPEILTHESLVGILQCLVAGPEPTVDVYFCGA 447
 QY 471 SPNCTEITE-GVMNRKANRKFQWVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILL 529
 DB 448 EQRCPPVPGPLDVQMGNSSVPSGKLVQSSIDYSAFKHNGTVECRASNNVGKT--SAFF 505
 QY 530 NSPQPPFFIQDN-----ISFYATIGVCLLFIWVLTLLCHIKYKKQFRYESQ 575
 DB 506 N----FAFKGNSKEQMHPTLFTPLLIGFVIAAGMMCIIVMILT----YKLOKPMYEVQ 557
 QY 576 LQMV-OVTGSSDNEYFYVDREYEDLKWEFPRENLEFGVLGSGAFKGYMNATAYGISK 634
 DB 558 WKVVEING---NNYVIIDTQPLPYDHKKWEPNRLSFGKTLGAGAFKGVVEATAYGLIK 614
 QY 635 TGVSIQVAVKMLKEKADSSREALMSLKMTOGSHENIVNLIGACTLSGPIYLIFEYC 694
 DB 615 SDAAMTVAVKMLKPSAHLTEREALMSLKVLSVLSGNHNMIVNLIGACTVGGPTLVITEYC 674
 QY 695 CYGDLNLYLRKREKFRHTWTEIFKEHNFSTYPT-FQSHPNSSMPGSE-VQIHPDSDQI 752
 DB 675 CYGDLNLFRRKRKDSFICSQEDHAE--VALYKNLLQSKESSCNDSTNEYMDMKPGYSYV 732
 QY 753 SGLHGSFSE---DEYENQKRLEEDLNVLTPEDLLCFAYQVAKGMEFLFEKSCVH 809
 DB 733 VPTKADRRSARGSTIERDVTPTAIMEDEL-ALDLELLSFSYQVAKGMAFLASKMCIH 791
 QY 810 RLDAARNLVTHGVKVKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGIYTIKS 869
 DB 792 RLDAARNILLTHGRITKICDFGLARDIKNSNVYKGNARLPVKWMAPESEFNCVYTFES 851
 QY 870 DWYSYGILLWEIFSLGNVPYGPVPDANFYKLQNGFKMDQPPYATETIYIMOSWAFD 929
 DB 852 DWYSYGIFLWELFSLGSPYPGMPVDISKYMIKEGFRMLSPHAPAEYMDIMKTCWDAD 911
 QY 930 SRKRPSFPNLTSLFGCOLADAEAEAYQNV 958
 DB 912 PLKRPTFKQIVQVLEIKQISDSTNHIYSNL 940
 RESULT 10
 KMS.FSVMD
 ID KMS.FSVMD STANDARD; PRT; 978 AA.
 AC P00545; Q86597;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).
 GN V-FMS.
 OS Feline sarcoma virus (strain McDonough).
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11778;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84119469; PubMed=6582485;
 RA Hampe A., Gobet M., Sherr C.J., Galibert F.;
 RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows
 RT unexpected homology with oncogenes encoding tyrosine-specific protein
 RT kinases";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).

QY 858 ESIFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIONGFKMDPEVATEE 917
 Db 857 ESIFDCVYTVQSDVWSYGILLWEIFSLGVNYPGILVNSKFKYKLVDGQYMAQAPAPKN 916
 QY 918 IYIIMQSWAFDRSRPSFNLTSFLGCLADAEAEAMYQNVDRV---SECPHTYQNR-- 972
 Db 917 IYVIMQACWALEPTRRPTQQICSLLOKQ---AQE-----DRRVPNTNLPSSSSRL 967
 QY 973 RPFSP 977
 Db 968 RPWOR 972

RESULT 11
 KEMS_HUMAN
 ID KEMS_HUMAN STANDARD; PRT; 972 AA.
 AC P07333;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Macrophage colony stimulating factor 1 receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms) (CD115 antigen).
 GN CSF1R OR FMS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89239490; PubMed=2524025;
 RA Hampe A., Shamoon B.M., Gobet M., Sherr C.J., Galibert F.;
 RT "Nucleotide sequence and structural organization of the human FMS
 RT proto-oncogene.";
 RL Oncogene Res. 4:9-17(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86175013; PubMed=2421165;
 RA Coussens L., van Beveren C., Smith D., Chen E., Mitchell R.L.,
 RA Isacke C.M., Verma I.M., Ullrich A.;
 RT "Structural alteration of viral homologue of receptor proto-oncogene
 RT fms at carboxyl terminus.";
 RL Nature 320:277-280(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97179223; PubMed=9027509;
 RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
 RA Hu W.X., Galibert F.;
 RT "Sequence analysis of two genomic regions containing the KIT and the
 RT FMS receptor tyrosine kinase genes.";
 RL Genomics 39:216-226(1997).
 RN [4]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=86261741; PubMed=2524648;
 RA Visvader J., Verma I.M.;
 RT "Differential transcription of exon 1 of the human c-fms gene in
 RT placental trophoblasts and monocytes.";
 RL Mol. Cell. Biol. 9:1336-1341(1989).
 RN [5]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=86281820; PubMed=3525854;
 RA Wheeler E.F., Roussel M.F., Hampe A., Walker M.H., Fried V.A.,
 RA Look A.T., Rettenmiller C.W., Sherr C.J.;
 RT "The amino-terminal domain of the v-fms oncogene product includes a
 RT functional signal peptide that directs synthesis of a transforming
 RT glycoprotein in the absence of feline leukemia virus gag sequences.";
 RL J. Virol. 59:224-233(1986).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- DATABASE: NAME=PROW; NOTE=PROW I:13-21(2000); 1999017666_g.htm".
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1999017666_g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M25786; AA58421.1; -
 CC EMBL: U63963; AA51696.1; -
 CC EMBL: M14002; AA35849.1; -
 CC EMBL: X03663; CAA27300.1; -
 CC PIR: S08123; TVHUMD.
 CC HSP: P11362; LFGL.
 CC Genew: HGNC:2433; CSF1R.
 CC MIM: 164770; -
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003598; Ig_C2.
 CC InterPro: IPR003600; Ig_Like.
 CC InterPro: IPR001824; RTKinaseIII.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00047; Ig; 4.
 CC Pfam: PF00069; pkinase; 1.
 CC ProDom: PD000001; Euk_pkinase; 2.
 CC SMART: SM00410; IG_Like; 3.
 CC SMART: SM00408; IGC2; 1.
 CC SMART: SM00219; TyrKc; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 CC Immunoglobulin domain; Repeat; Polymorphism.
 FT SIGNAL 1 19
 FT CHAIN 20 972
 FT MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR.
 FT DOMAIN 20 512
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 513 537
 FT POTENTIAL.
 FT DOMAIN 538 972
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 24 104
 FT IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 107 197
 FT IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 204 298
 FT IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 299 399
 FT IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 400 505
 FT IG-LIKE C2-TYPE DOMAIN 5.
 FT DOMAIN 582 910
 FT PROTEIN KINASE.
 FT NP_BIND 588 596
 FT ATP (BY SIMILARITY).
 FT BINDING 616 616
 FT ATP (BY SIMILARITY).
 FT ACT_SITE 778 778
 FT DISULFID 42 84
 FT DISULFID 127 177
 FT POTENTIAL.
 FT DISULFID 224 278
 FT POTENTIAL.
 FT DISULFID 419 485
 FT POTENTIAL.
 FT MOD_RES 699 699
 FT PHOSPHORYLATION (AUTO-) (IN VITRO) (BY
 FT SIMILARITY).
 FT MOD_RES 708 708
 FT PHOSPHORYLATION (AUTO-) (IN VIVO) (BY
 FT SIMILARITY).
 FT MOD_RES 809 809
 FT PHOSPHORYLATION (AUTO-) (IN VITRO) (BY
 FT SIMILARITY).
 FT CARBOHYD 45 45
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 73 73
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 153 153
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 240 240
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 275 275
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 302 302
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 335 335
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 353 353
 FT N-LINKED (GLCNAC...) (POTENTIAL).

```

Db      835 QSDVMSYGILLWEITFSLGNLPYGLVNSKYKLVKGYQMAQFAFAFKNTIYSIMQACWA 894
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy      928 FDSRFSFPLNTSLFLGCOL-ADAEAWYQNV 958
      :||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      895 LEPHTRPTFOICSLFQEQEDRRERYTNL 926
      :||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
KIT_CHICK
ID KIT_CHICK STANDARD; PRT; 960 AA.
AC Q08156;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.II2) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (C-kit).
GN KIT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Ieghorn; TISSUE=Brain;
RX MEDLINE=9329295; PubMed=7685729;
RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,
RA Naito M., Sakurai M.;
RT "Cloning and expression of the chicken c-kit proto-oncogene.";
RL Gene 128:257-261(1993).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO
CC PRESENT IN THE BURSA OF FABRICIUS, HEART, KIDNEY, LUNG, SPLEEN
CC THYMUS AND OVARY.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13225; BAA02506.1;
CC PIR; JN0677; JN0677.
CC HSSP; P11362; IFCG.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR001824; TyrKinaseIII.
CC InterPro; IPR001125; Tyr_pkinase.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.
CC PRODOM; PD000001; Euk_pkinase; 2.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
CC Immunoglobulin domain.

```

FT SIGNAL 1 24
FT CHAIN 25 960
FT DOMAIN 25 505
FT TRANSMEM 506 530
FT DOMAIN 531 960
FT DOMAIN 531 960
FT DOMAIN 575 913
FT NP_BIND 581 589
FT BINDING 609 609
FT ACT_SITE 777 777
FT MOD_RES 808 808
FT CARBOHYD 76 76
FT CARBOHYD 135 135
FT CARBOHYD 149 149
FT CARBOHYD 269 269
FT CARBOHYD 286 286
FT CARBOHYD 306 306
FT CARBOHYD 318 318
FT CARBOHYD 338 338
FT CARBOHYD 343 343
FT CARBOHYD 356 356
FT CARBOHYD 453 453
FT CARBOHYD 469 469
SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;

Query Match 23.0%; Score 1215; DB 1; Length 960;
Best Local Similarity 32.3%; Pred. No. 9.6e-74;
Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;

QY 66 ALRPOSSCTVYAAAVEVDVSAITLQVLVDAPGNISLWVFKHSS-----LNCOPHFD 119
DB 18 SLIPAGGVPHEESSLVNKGELKNEGP-----VTWNQNSDPSAKTRISNEKEWH 73
QY 120 LONRGV--VSMVILKMTQAGYLLFIOSEATNVTILFTVIRNLTLLRYPYRKE 177
DB 74 TKNATIRDIGRECKSGSIVNSVYVFKDP-----NVLFV--DSLIIY-----GKED 119
QY 178 NODALVCSIESPEP-IVEWLDCSQGESCKEESPAVVKKE-----KVLHELFGTDIRCC 232
DB 120 SDILLVC---PLTDPDLNFTLRKCDGRLPKNMTFIPNPKQGIIRKIVORSFKGCYQCL 176
QY 233 AR-NELGRECTRLFLIDNQT-P-QTTLPOL-----FLKVGGEPLWIRKAVHNVHGFGL 283
DB 177 AKHNGVEKISEHIF---LNVRPVHKALPVITLSKSYELLKEGEEPEVTCIITDVS SVKA 233
QY 284 TWELNKALEEGNYFEMSTYNTNTRMIRLFAVSVSARNDTGYTCSSKHP--SOSAL 341
DB 234 SWISYKSAIVTSKRNGLDGYERK---LTLNIRSGVNDSGEFTC-QAENPFGKTNAT 288
QY 342 VTI--VGKGFN---ATNSSEDEYDQYEEFCFSVRFKAYPQIR-CTWTFSRKSPFCEQK 395
DB 289 VTLKALAKGFVRLFATMTNTTIDINAGQNGN--LTVEYEAYPKPEEVMYMNTEL---QN 343
QY 396 GLDNGYSISKFNKHH-----OPGEYIPHAENDDAQFTKMTLINIRKPOVLA 443
DB 344 SDDHYVKTKVGNNSYTSSELHLTRLKGTEGGIYTFVFSNSDASSSVTFNVYVYKTEILT 403
QY 444 EASASQ--ASCFSDGYPLSPWTKKCDKSPNCTEBITEGVWNRKANRKVFGQWVSSTL 501
DB 404 LDMGLNDILQCVATGPAPTIVYFCPGTEQRCCLDSPTISPMDBKVS-----YTNSVSP 457
QY 502 NMSEAIKGLVKCCAYNSLGTSCETILLNSPGFPF-----IQDNISFYA-----TIGV 550
DB 458 SPERILVESTVNASMFKSTGTICCEASSNGDKSSVFNFFAIKEQIRHTLFTPLLLAFJG 517
QY 551 CLLFIWLTLLCHKKYKFRYESQLQW--QVTGSSDNEFYVDFEYEDLKWEPFREN 609
DB 518 AGLMCIIVWILYIYLPKPKVEQVQVVEING---NNYVIDPTQLPVDHKEWEPFNR 574
QY 610 LFFGKVLGSGAFGKVNATAYGISGTQVAVKMLKEKADSSEREAALMSLKMQLTG 669
DB 575 LSGFKTLGAGKGVVEATAYGLFKSDAAMTAVKMLKPSAHLTEREALMSLKLVSILG 634
QY 670 SHENIVNLLGACTLSGPIYILIFEYCCYGDLLNLYRSKREK-----HRTWTEIFKEHNFSF 725

DB 635 NHINIVNLLGACTIGPTLVITEYCCYGDLLNLRKRDSFCPKHEEHAEEAAYENL-- 692
QY 726 YPTFOSHPS-----SMPGSRVOIHDPDSOISGLHNSFHSDEIEIYENOKRLEEE 778
DB 693 --LHOAEPTADAVNEYMDKPGSVYAVPPKADKKRVKSGSYTDOD-----VTLSMLEDD 746
QY 779 DLNVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDGLGARDIMS 838
DB 747 --LALDVELLSFSQVQAGMSFLASKNCIHRDLAARNLLTHRTKICDGLGARDIRN 804
QY 839 DSNVYVGNARLPVKWMAPESEFEGYITIKSDVYSYGILLWEIFSLGVNPPYGPVDANF 898
DB 805 DSNVYVGNARLPVKWMAPESEFEGYITIKSDVYSYGILLWEIFSLGVNPPYGPVDANF 864
QY 899 YKLQNGFKMDQFPYATEEIIYIMQSWAFDSKRSPNLTSLFGCOLADAEAMYQNV 958
DB 865 YKMIKEGYRMFSPECSPPMEYDINKSCWDADLPQRTFKQIVOLIEQOLSDNAPRVYAN- 923
QY 959 DGRVSECPHYQN 971
DB 924 ---FSTPPSTQGN 933

RESULT 13
KPM5_MOUSE STANDARD; PRT; 977 AA.
AC P09581; Q9DBH9;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR CSFMR OR FMS.
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88217329; PubMed=2966922;
RA Rothwell V.M., Rohrschneider L.R.;
RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
RT expression.";
RL Oncogene Res. 1:311-324(1987).
RN [2]
RP Rothwell V.M.;
RA Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RX MEDLINE=93181280; PubMed=8441691;
RA de Parseval N., Boreaux D., Gisselbrecht S., Sola B.;
RT "Reassessment of the murine c-fms proto-oncogene sequence.";
RL Nucleic Acids Res. 21:750-750(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."; [5]
 RL Nature 409:685-690(2001).
 RN
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=93268269; PubMed=8497248;
 RA Yue X., Favot P., Dunn T.L., Cassidy A.I., Hume D.A.;
 RT "Expression of mRNA encoding the macrophage colony-stimulating factor
 RT receptor (c-fms) is controlled by a constitutive promoter and tissue-
 RT specific transcription elongation."; [5]
 RL Mol. Cell. Biol. 13:3191-3201(1993).
 RN [6]
 RP AUTOPHOSPHORYLATION SITES.
 RX MEDLINE=90258890; PubMed=2160591;
 RA van der Geer P., Hunter T.;
 RT "Identification of tyrosine 706 in the kinase insert as the major
 RT colony-stimulating factor 1 (CSF-1)-stimulated autophosphorylation
 RT site in the CSF-1 receptor in a murine macrophage cell line."; [5]
 RL Mol. Cell. Biol. 10:2991-3002(1990).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X06368; CA229666.1; ALT_SEQ.
 DR EMBL; AK004947; BAB23691.1; -.
 DR EMBL; S62219; -; NOT_ANNOTATED_CDS.
 DR PIR; S01880; TVMSMD.
 DR HSSP; P11362; 1FGK.
 DR MGD; MGI:1339758; Csf1r.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR001824; RtkinaseII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 4.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00410; IG_Like; 3.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 19
 FT CHAIN 20 977
 FT MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR
 FT EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 20 511
 FT TRANSMEM 512 536
 FT DOMAIN 537 977
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 24 104
 FT IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 107 197
 FT IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 204 298
 FT IG-LIKE C2-TYPE DOMAIN 3.

FT	DOMAIN	299	397	IG-LIKE C2-TYPE DOMAIN 4.
FT	DOMAIN	398	503	IG-LIKE C2-TYPE DOMAIN 5.
FT	DOMAIN	580	913	PROTEIN KINASE.
FT	NP_BIND	586	594	ATP (BY SIMILARITY).
FT	BINDING	614	614	ATP (BY SIMILARITY).
FT	ACT_SITE	776	776	BY SIMILARITY.
FT	DISULFID	42	84	POTENTIAL.
FT	DISULFID	127	177	POTENTIAL.
FT	DISULFID	224	278	POTENTIAL.
FT	DISULFID	417	483	POTENTIAL.
FT	MOD_RES	697	697	PHOSPHORYLATION (AUTO-) (IN VITRO).
FT	MOD_RES	706	706	PHOSPHORYLATION (AUTO-) (IN VIVO).
FT	MOD_RES	807	807	PHOSPHORYLATION (AUTO-) (IN VITRO).
FT	CARBOHYD	45	45	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	73	73	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	335	335	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	389	389	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	410	410	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	449	449	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	478	478	N-LINKED (GLCNAC) (POTENTIAL).
FT	CARBOHYD	491	491	N-LINKED (GLCNAC) (POTENTIAL).
FT	CONFLICT	446	446	Q -> E (IN REF. 1).
FT	CONFLICT	553	553	R -> S (IN REF. 1).
FT	CONFLICT	616	616	L -> I (IN REF. 1).
FT	CONFLICT	744	744	A -> H (IN REF. 1).
FT	CONFLICT	814	814	MISSING (IN REF. 1).
SEQ	SEQUENCE	977 AA;	109178 MW;	7EDF8310CCCF98906 CRC64;

Query Match 22.8%; Score 1204; DB 1; Length 977;
 Best Local Similarity 33.2%; Pred. No. 5 3e-73;
 Matches 322; Conservative 155; Mismatches 330; Indels 164; Gaps 32;

QY	83	VDSASITLQVL----	VDAPGNISCLWVFKHSSLCNCPHFDLQNRGVV-----	SNVLKLM	133
DB	32	VEPGETVTLRCVSGSVEMDGPISPYWTLDPES----	PGSLTTRNATFKNTGYRCTEL	87	
QY	134	TEQAGELYL--LFIOSEATNYTIL---	FTYSIRNTLYTLRRPYFRKMNEDALVCISES	188	
DB	88	EDPMAGSTTHLYVKDPAHWNLLAQEVTV-----	-----VEGEAV-----	124	
QY	189	VPEPIVEMVLCDSOGESCKEESPAAVKK-----	-----EBKVLHELFGTDIRCCAR	234	
DB	125	LPCLITDPALKDSV-SLMREGGROVLKTYVFFSPWRGFIIRKAKVL---	-----DSNTYVCKT	179	
QY	235	NELGRECTRL-FTIDLN-----	OTPQTLT-POLFLKV-GEPLWIRKAVHNVHGLTWEL	287	
DB	180	MVNGRESTGIVLKVNRVHPEPPOIKLEPSKLVIRGEAAQIVCSATNAEVEGVILKR	239		
QY	288	ENKALE-----	EGNYFEMSTYSTNTMTIRILPAFVSSVARNDTGYTC-SSSKHPSQS	339	
DB	240	GDTKLEIPLNSDFQDNYYK-----	-----KVRALSLNAVDFOAGIYSCVASNDVGTRT	288	
QY	340	ALVT--IVGKGFINATN--SSEDIYDIOVEEFCFSVRKAYPOIR-CTWTFSRKSPFCEOK	395		
DB	289	ATMNFQVVEASAYNLNLTSEQLQEVSGDSLILTVHADAYPSIOHYNWYTLGPFEF-DQR	347		
QY	396	GLD-----	NGYSISKFCNH--KHOPGEYIFHAENDDAQFTKMTFLNIRKQPVLAES	446	
DB	348	KLEFITORAIRYRTFKLFLNRVKASEAGQYFLMAONKAGWNNTLFTLRLRYPPVSVTW	407		
QY	447	ASQAS-----	CFSDGYPLPSWTWKCKSDKSPNCTEITEEGVWN-----RKANKRVGQWVSS	498	
DB	408	PVNGSDYLFCDVSGYQPSVTWMECRGHTDRCDQAQLOVWMDTHPEVLSOKPDKVLIQ	467		
QY	499	STLNMSIAKGLFKCCAYNSLGTSCETILLNSPGPFPIODNISFYATIGVC---	LLFI	555	
DB	468	SOLPIGTLKINMTYFCKTHNSVGNSSQYFRAVLSQSKQLPDESFLTTPVVVACMSVMSLL	527		
QY	556	VVLTLTLLCHKYKKOFRYESQLQWQVGTGSSDNEYFYVDREYEDLKWEPFRENLEFGK	615		
DB	528	VLLLLLLLLLYKKQPKYQVRWKIIE--RYEGNSYTFIDPTOLPYNEKWEFFRNNOFGKT	585		

QY 616 LGSGAGKVMNATAYGISTGYSIOVAVKMLKEKADSSERREALMSKMTOLGSHENIV 675
 Db 586 LGAGAFGKVVATEAGLGKEDAVLVAVKMLASTAHADREALMSKMTOLGSHENIV 645
 QY 676 NLGACTLGGPIVLYFEYCCYGDLLNLYLSKRE-----KFRHTWT 715
 Db 646 NLGACTHGGPVLVITEYCCYGDLLNLYLSKRE-----KFRHTWT 715
 QY 716 EFKKEHNSFYPTFOSHNSMPGSGREVOIHDPDSOISGLHNSHSEDEIEYENOKRLE 775
 Db 706 YVRDSGFS-----SQGVDTYVEMRPVSTSSDSFEKQD-LDKEASRPLE 749
 QY 776 BEEDNLVTFEDLLCFAYOVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARD 835
 Db 750 -----LMDLLHFSVOAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800
 QY 836 IMSDSNVYVRGNARLPVKWMAPESEFEGYTIKSDVMSYGILLWEIFSLGVNPPGIPVD 895
 Db 801 IMSDSNVYVRGNARLPVKWMAPESEFEGYTIKSDVMSYGILLWEIFSLGVNPPGILVN 860
 QY 896 ANFYKLIONGFKMDOPFYATEEIIYIMOSCAFDSSKRSFNLTSFLGCOLADAEMY 955
 Db 861 NKFXLVKVDGYMAQGVFAKPNYISIMOSCDLEPTTRPTFOQICFL-----QEOARL 914
 QY 956 QNVGVRSECP 966
 Db 915 ERRDQDYANLP 925

RESULT 14
 KFMS_RAT
 ID KFMS_RAT STANDARD; PRT; 978 AA.
 AC Q00495;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
 GN CSF1R OR CSFMR OR FMS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID-10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Wistar; TissUP-Skeletal muscle;
 RX MEDLINE-93001225; PubMed-1389227;
 RA Borycki A.G.; Guiller M.; Leibovitch M.P.; Leibovitch S.A.;
 RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence
 analysis and regulation during myogenesis.";
 RL Growth Factors 6:209-218(1992).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X61479; CAA43706.1;
 DR PIR; S16385; S16385.
 DR HSP; P11362; IFGK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR003600; Ig_like.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 4.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR SMART; SM00410; IG-like; 3.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; IYK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 19
 FT CHAIN 20 978
 FT BY SIMILARITY.
 FT MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE DOMAIN 1.
 FT IG-LIKE C2-TYPE DOMAIN 2.
 FT IG-LIKE C2-TYPE DOMAIN 3.
 FT IG-LIKE C2-TYPE DOMAIN 4.
 FT IG-LIKE C2-TYPE DOMAIN 5.
 FT PROTEIN KINASE.
 FT DOMAIN 20 511
 FT TRANSMEM 512 536
 FT DOMAIN 537 978
 FT DOMAIN 24 104
 FT DOMAIN 107 197
 FT DOMAIN 204 298
 FT DOMAIN 299 397
 FT DOMAIN 398 503
 FT DOMAIN 580 914
 FT NP_BIND 586 594
 FT BINDING 614 614
 FT ACT_SITE 776 776
 FT DISULFID 42 84
 FT DISULFID 127 177
 FT DISULFID 224 278
 FT DISULFID 417 483
 FT MOD_RES 697 697
 FT MOD_RES 706 706
 FT MOD_RES 807 807
 FT CARBOHYD 45 45
 FT CARBOHYD 73 73
 FT CARBOHYD 302 302
 FT CARBOHYD 335 335
 FT CARBOHYD 389 389
 FT CARBOHYD 410 410
 FT CARBOHYD 449 449
 FT CARBOHYD 478 478
 FT CARBOHYD 491 491
 SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;
 Query Match 22.6%; Score 1194; DB 1; Length 978;
 Best Local Similarity 33.1%; Pred. No. 2.5e-72;
 Matches 321; Conservative 154; Mismatches 332; Indels 164; Gaps 32;
 QY 83 VDSASTLQVL-----VDAPGNISCLWVFKHSSLCNCPHFDLQNRGV-----SWVILKM 133
 Db 32 VEPGETVTLRCVSGSVEWDGPISPYWTLDPS-----PGSTLTTRNATFKNTGYRCTEL 87
 QY 134 TETQAGEYL-LFLOSATNYTL---FTYSIRNTLLYLRPFYRKMNODALVCISES 188
 Db 88 EDPMAGSTTHLYVKDFAHNSWNLQAQEVTV-----VEGEAV----- 124
 QY 189 VPEIVEWLCLDSOGESCKESPAAVKK-----EKKVHELFEFGTDIRCCAR 234
 Db 125 LPCLITPAKDSV-SLMREGROVLKTYVFFSAWRCFIIRKAKVL-----DSNTYVCKT 179
 QY 235 NELGRECTRL-FTIDLN-----QTPOTTL-PQLFLKV-GEPLWIRCKAVHNVHGLTWEL 287
 Db 180 MVNGRESTSTGIWLKVNVRVHPPEPOIKLEPSKLVIRGEAAQIVCSATNAEVEGFNVILKR 239

```
QY 288 ENKALE-----EGNYFEMSTYSTRNTRILFAFVSSVARNDTGYTC-SSSKHPSQS 339
Db 240 GDTKLEIPLNSDFQDNYK-----KYRALSINAVDFQDAGIYSCVANDVGTRT 288
QY 340 ALVT--IVQKGINATN--SSEYDEIDQYEEFCFSVFKAYPQIR-CTWTFSRKSFCEQK 395
Db 289 ATMFQVVEISAYLNTSEQLLQVEVSGSLITLVHADAYPSIQHNWTVLGVFFE-DQR 347
QY 396 GLD-----NGYSISKFNH--KHOPGYIIFHAENDDAQFTKMFILNRRKPOVLAEAS 446
Db 348 KLEFTQRIAYITFKLFILNRVYKASEAGYFLMAONKAGWNNTFFELTRYPEVSVTW 407
QY 447 ASQAS-----CFSDGYPLPSWTWKCKSDKSPNCTEETEGVNN-----RKANRKVFGOWSS 498
Db 408 PYNGSDVLFCDVSGVPQPSVTWMECRGHDRDCDEAALQVNDTHPEVLSQRPFKVII 467
QY 499 STLNSEAIKGLVKKCAVNSLGTSETILLNSPGPFPIQDNISYATIGVC-----LLFI 555
Db 468 SOLPIGTLKHNNITYFCKTHNSVGNSSQYFRAVSLGQSKQLPDBESLFTPVVACMSVMSLL 527
QY 556 VVLTLLCHIKYKQFYESQLOLQVQVGTSGSDNEFYFVDFREYEDLKWEPRENLEFGK 615
Db 528 VILLLLLYKKYKOKPKYQVRWKIIE--RYEGNSYTFIDPTQLPYNEKWEFFPRNLQFGKT 585
QY 616 LGSAGFGKVMNATAYGISKTVGSIOAVKMLKEKADSSEREALMSLKMNTOLGSHENIV 675
Db 586 LGAGAFGKVEATEAGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMNTOLGSHENIV 645
QY 676 NLLGACTLSGPIYLIFCYCYGDLNLYLSKRE-----KFHRTWT 715
Db 646 NLLGACTHGGPVLVITEYCCYCGDLLNFLRKAEALMGLSPQDSEGDSSYKNIHLEKK 705
QY 716 EFKFHNFSFYTFQSHPNNSMPGSEVOIHDPDSOISLGHGNSFSEDEIEYENOKRLE 775
Db 706 VYRDSGFS-----SQGVDTYVEMRPVSTSSDSDFEKQD-LDKEPSRPLE 749
QY 776 BEEDNLVTFEDLLCFAYOVAKGMEFLKSCVHRDLAARNVLVTHGVKVKICDFGLARD 835
Db 750 -----LWDLHFSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800
QY 836 IMDSNVYVGNARLPVKWAPESIFEGYITKSDVMSYGILLWEIFSLGVNPPYGPVD 895
Db 801 IMDSNVYVGNARLPVKWAPESILCYVTVQSDVMSYGILLWEIFSLGVNPPYGPILVN 860
QY 896 ANFYKLIQNGFKMDQFPYATEIYIIMQSCWAFDSCRKRSFPLNLTSLFCQLADAEAMY 955
Db 861 NKFKYLVKDGQYMAQVPEAFKPIYSIMQSCWDLPTRRPTFOQICFLL-----QEQARL 914
QY 956 QNVGGRVSECP 966
Db 915 ERRODYANLP 925

RESULT 15
PGDS_RAT
ID PGDS_RAT STANDARD; PRT: 1088 AA.
AC P20786;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha platelet-derived growth factor receptor precursor (EC 2.7.1.112)
DE (PDGF-R-alpha).
GN PDGFRA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=90220609; PubMed=2157969;
RA Lee K.H., Bowen-Pope D.F., Reed R.R.;
RT "Isolation and characterization of the alpha platelet-derived growth
```

```
RT factor receptor from rat olfactory epithelium.";
RL Mol. Cell. Biol. 10:2237-2246(1990).
RN [2]
RP SEQUENCE OF 33-524 FROM N.A.
RX MEDLINE=93305723; PubMed=8318839;
RA Herren B., Weyer K.A., Rouge M., Loetscher P., Pech M.;
RT "Conservation in sequence and affinity of human and rodent PDGF
RT ligands and receptors.";
RL Biochim. Biophys. Acta 1173:294-302(1993).
CC -!- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M63837; AAA40743.1; ALT_INIT.
CC PIR: A34710; PERTCA.
CC HSSP: P11362; IFGK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; Ig; 4.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 2.
CC SMART: SM00409; IG; 3.
CC SMART: SM00219; Tyrc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein;
CC Transferase; Phosphorylation; ATP-binding; Immunoglobulin domain;
CC Signal.
CC SIGNAL 1 23
CC CHAIN 24 1088
CC
CC ALPHA PLATELET-DERIVED GROWTH FACTOR
CC RECEPTOR.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC PROTEIN KINASE.
CC ATP (BY SIMILARITY).
CC ATP (BY SIMILARITY).
CC BY SIMILARITY.
CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC O-LINKED (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC L -> R (IN REF. 2).
CC S -> T (IN REF. 2).
CC CONFLICT 519 519
CC SEQUENCE 1088 AA; 122641 MW; 590C8BB0418601E7 CRC64;
```

```
Query Match      22.3%; Score 1176; DB 1; Length 1088;
Best Local Similarity 35.6%; Pred. No. 4.6e-71;
Matches 294; Conservative 128; Mismatches 288; Indels 115; Gaps 24;

Qy 246 TIDLNQTPOTTLPQLFLKVGELWIRKXAVHVNHGFLTW-----ELENKA-----LEEGNYF 298
Db 210 TSELNLEMDTR--QTVYKAGETIVVTC-AVFNEVVDLQWTPYGEVRNKGITMLEE---- 262

Qy 299 EMSYSTNRTWIRILFAP-VSSVARNDGYTCSS--SSKHPSQSALVTIV--GKGFIN- 351
Db 263 -----IKLPSIKLYVITVTKATVKDGYECAARQATKEVKEMKTVTISVHERGFOI 316

Qy 352 --ATNSSDEYEDQYEEFCFSVRFRKAYPOICRTWTFSRKSF-----PCBQKGLDNGY 401
Db 317 RPTFGHLETVNLHQVREFV--VEVQAYPTPRISWLKDNLTLIENLTEITTDVQORSQETRY 374

Qy 402 -SISKFCNHHQO-PGEYIFHAENDDAQFTKMETLINIRKQVVL-----AEASASQASC 452
Db 375 QSKLLIRAKEDSGHYTIIQNDDDMKSYTFELSTLVPASILELVDDHHGSGGGGTVC 434

Qy 453 FSDGYPLPSWTH-----KKCSDK-----SPNCTEETEGVWNRKANKRVFGOWVSSST 500
Db 435 TAEGTPLPNIEMICKDKKNDTSWTLASVNSNIITE--FHGRGSTVEGR----- 486

Qy 501 LNMSEAIKGFVKCCAYNSLGTSCETILLNSPGPPFIQDNISFYATICGICLLFIVVLT 560
Db 487 VSAKVEETIARCLAKNDLIGNRELKLVAPS----LRSELTVAARVLLVIVIVSLI 542

Qy 561 LICHYKKQFRYEOSLOWQVVTGSSDNEYVYDFREYEDLKWEPPRENLEFGVLGSGA 620
Db 543 VLVVIWKQKPRYIWRVIESISPDGHEVIYVDPQMPYDSRWEPRDGLVLGRLGSGA 602

Qy 621 FGVVNATAYGISKTVGSIQVAVKMLKEKADSERALMSELKMMTQLGSHENIVNLLGA 680
Db 603 FGVVVEGTAYGLSRQPVKAVKMLKPTARSSEKQALMSELKIMTHLGPHLNIVNLLGA 662

Qy 681 CTLGSPVILIFCYCCYGLLNLVLSKREKF-----HRTWTEI 717
Db 663 CTKSGPIVITTEYCFYGDVNLVHLKNRDSFMRHPEKPKKDDIFGLNPADESTRSIVIL 722

Qy 718 FKEHNFSEYPTFQSHPNSSMPGSRVQIHPDSQISGLHGN--SPHSEDEIEYENQKRL 775
Db 723 SPENNGDYVDMKQADTTQVPMLEKREKVSYSQIQRSLYDRPASYYKKKSMLDSE-AKNLL 781

Qy 776 EEDLNLVTFEDLLCFAYOVAGMGEFLEPKSCVHRDLAARNVLVTHGVKVCDFGLARD 835
Db 782 SDDDSSEGLTLLDLSFTYQVARGMEFLASKNCVHRDLAARNVLLAQQKIVKICDFGLARD 841

Qy 836 IMSDSNYVVRGNARLPVKWMAPELFEGLYTTKSDVMSYGIILLWEIFSLGVNYPYGPV 895
Db 842 IMHDSNYVSKGSTFLPVKWMAPESIFDNLTYTLLSDVMSYGVLLWEIFSLGTPYGMVMD 901

Qy 896 ANFYKLIQNGFRMDOPFYATEEYIIMQSCWAFDSRKRFPFPNLTSLFGCQLADAEAMY 955
Db 902 STFYNNKIKSGYRMAKPDHATSEVYEIMVOCWNSPEKRPFSYHLSEIVENLLPGQYKKS 961

Qy 956 QNV--DGRVSECPH-----TYQNRPRFSSREMDLGL 983
Db 962 EKIHLDLFLKSDHPAVARMVRVSDNAYIGVTKYNEEDKLEWEGGL 1006
```

Search completed: May 27, 2003, 14:32:21
Job time : 29.0091 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:26:24 ; Search time 71.0358 Seconds

(without alignments)
2880.307 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*
- 15: sp-rviro:*
- 16: sp-bacteriap:*
- 17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1271.5	24.1	976	13 Q9W755	Q9W755 brachydanio
2	1265.5	24.0	974	11 Q63702	Q63702 rattus ratt
3	1263.5	24.0	978	11 Q63116	Q63116 rattus norv
4	1262.5	23.9	964	6 Q97744	Q97744 sus scrofa
5	1261.5	23.9	964	6 Q97Q01	Q97Q01 sus scrofa
6	1261.5	23.9	964	6 Q97Q00	Q97Q00 sus scrofa
7	1250.5	23.7	979	6 Q8WN23	Q8WN23 canis fami
8	1247	23.6	972	4 Q99662	Q99662 homo sapien
9	1244	23.6	978	6 Q9XS93	Q9XS93 canis fami
10	1230.5	23.3	977	13 Q98S01	Q98S01 danio nigro
11	1229	23.3	945	6 Q77589	Q77589 equus cabal
12	1228	23.3	974	13 Q98S03	Q98S03 danio dangi
13	1226	23.2	954	13 Q91909	Q91909 xenopus lae
14	1223.5	23.2	977	13 Q918N6	Q918N6 brachydanio
15	1221	23.2	948	6 Q9TTD7	Q9TTD7 trichosurus
16	1219	23.1	724	6 Q9MYN0	Q9MYN0 bos taurus

ALIGNMENTS

RESULT 1

Q9W755	PRELIMINARY;	PRT;	976 AA.
ID	Q9W755		
AC	Q9W755;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Kit receptor tyrosine kinase.		
GN	KIT.		
OS	Brachydanio rerio (Zebrafish) (Zebra danio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99396707; PubMed=10393121;		
RA	Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;		
RT	"Zebrafish sparse corresponds to an orthologue of c-kit and is		
RT	required for the morphogenesis of a subpopulation of melanocytes, but		
RT	is not essential for hematopoiesis or primordial germ cell		
RT	development."		
RL	Development 126:3425-3436(1999).		
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-		
CC	PROTEIN KINASES.		
DR	EMBL: AF153446; AAD41890.1; -		
DR	HSSP: P11362; IFGK.		
DR	2FIN; ZDB-GENE-980526-464; kit.		
DR	InterPro: IPR000719; Euk_pkinase.		
DR	InterPro: IPR003598; Ig_C2.		
DR	InterPro: IPR003600; Ig_Like.		
DR	InterPro: IPR003006; Ig_MHC.		
DR	InterPro: IPR001824; RTKinaseIII.		
DR	InterPro: IPR001245; Tyr_pkinase.		
DR	Pfam: PF00047; ig; 3.		
DR	Pfam: PF00069; pkinase; 1.		
DR	ProDom: PD000001; Euk_pkinase; 2.		
DR	SMART: SM00408; Igc2; 1.		

Q98SU2 danio kerri
P79750 fugu rubrip
Q98SU4 danio albol
Q9DE49 brachydanio
Q9PUF6 gallus gall
Q8R406 rattus norv
P79749 fugu rubrip
Q8UV8 fugu rubrip
Q97745 sus scrofa
Q8UV9 fugu rubrip
Q925F7 rattus norv
Q9PT10 brachydanio
Q8UW9 brachydanio
Q8QH13 gallus gall
Q8VCD0 mus musculu
P79701 coturnix co
Q9EQ22 rattus norv
Q91Z1 rattus norv
Q9EQ24 rattus norv
Q96KM0 homo sapien
Q96KM2 homo sapien
Q96KM1 homo sapien
Q90749 gallus gall
Q96KL9 homo sapien
Q9TT23 oryctolagus
Q91416 xenopus lae
Q91286 pleurodeles
Q14672 homo sapien
Q91287 pleurodeles

```
DR SMART: SM00410; IG-like; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 976 AA; 109278 MW; A80AA01658C1A2DA CRC64;

Query Match 24.1%; Score 1271.5; DB 13; Length 976;
Best Local Similarity 38.5%; Pred. No. 3.4e-96;
Matches 294; Conservative 122; Mismatches 262; Indels 85; Gaps 20;

QY 253 POTTLPQ---LFLKVGEPILWIRKAVHVNHGFLTWLE---NKALEEGNYFEMSYSTN 306
DB 206 PPITLGPQORVLLTQGEKLSLGSSTNSVNSDIAVKWKPAGNPVSVHNSHLLETPITHV 265
QY 307 RTMIRTLFAFVSSVARNDTGYYTCSSSKHPSQA--LVTIVGKGFNATN--SEDEYID 362
DB 266 RTAILLSL---SSVTQMDAGNICEAINEKGTAKPVMWNIYEKGFNITSDVNSRPRV 321
QY 363 QVEEFCFSVRKAYPQIRCTWTFSRKSPCEQKLDNGYSISKFCNH----- 409
DB 322 AGESLSLRVVMNAYPKPHFSWSYS-----GVKLTNTDHTVITSRTHGNSYT 368
QY 410 -----KHQGEYIFHAENDDAQFTKMTLNIRKPKQVLAESA--SQASCFSQDGP 458
DB 369 SELKLVRLKVSIGIVTFCSLNRDATIRQTFEVHISKPOIVSYEGIDQGVRCVABGYP 428
QY 459 LPSWTW-----KCKSDKSPNCTEETEGVWNRKANRKFVGOWVSSSTNSMSEAIGKFL 511
DB 429 TPOIKYICDLPHSRCSNLL-NATQE-EEDVVTWNTNPPFGKAVESRLNITK--NNYA 484
QY 512 VKCAYNSLCTSETILLNSPGPFP---FIQDNISFYATIGCVLLFIVLTLTCHRYKK 568
DB 485 TLECVASANGEIVYTLFISENTVPHELTPLLIGFVAAAVILVILVLT---YKYM 540
QY 569 QFYRESOLQWVQSSDNEYFYVDREYEDYDLKWEPPRENLEFGKVLGSGAFKGMVNT 628
DB 541 KPKYQIQWKVIE--GIHGNVNVYIDPTQLPYDHOWEFPKRLFGKTLGSGAFKGVYEA 598
QY 629 AYGSKTGVSIQVAVKMEKADSSEREAALMSKMMTOLGSHENIVNLLGACFTLSGPY 688
DB 599 AYGMSKADTVMTAVKMLKPSAHATEKALMSKLVLSYLGHNHINIVNLLGACVGGPTL 658
QY 689 LIFCYCCYGLLNLRLSKRKFHRTWTEIFKEHNFSEYPTFQSHPNSSMPGSRVQ---- 744
DB 659 VITEYCCFGDLNLFRRRVVYFYT---TLGEDAYRNVMMQSEPNDSRNGYMTMKPSVL 715
QY 745 -IHPDSQISGLHGNSTHSEDEIYENQKRLLEEEDNLVLTFDLLCFAYQVAKGMEFLE 803
DB 716 GILLSERNRSLNKGDYSYSDAV-----SEILOEDGLTLDT-EDLLSFSYQVAKGMDFLA 769
QY 804 FKSCVHRDLAARNLVTHGKVKVYKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEG 863
DB 770 SKNCIHRDLAARNLLTQGRVAKICDFGLARDITDSNVYVKGARLPVKWMSPESTFEC 829
QY 864 IYTKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGKFMQDPFYATEEIIYIMQ 923
DB 830 VYTFESDWSYGILLWEIFSLGSSPYPCMPVDSKPYKMIKEGYRMESPEFSPEMYDIMH 889
QY 924 SCWAFDSKRKPSFNLNLSFLGCOLADAEAMQVNDGRVSECP 966
DB 890 SCWDADVPKPSFSKIVEIKBQIQSDSTKHIIYLFSSRLPAAP 932

RESULT 2
Q63702 PRELIMINARY; PRT; 974 AA.
ID Q63702
AC Q63702;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
```

```
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE C-kit receptor tyrosine kinase isoform.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN/FMAI;
RA Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;
RT "Two isoforms of rat c-kit receptor tyrosine kinase.";
RL Nucleic Acids Res. 0:0-0(0).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; X62491; CAA44354.1; -
DR HSSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_Like; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;

Query Match 24.0%; Score 1265.5; DB 11; Length 974;
Best Local Similarity 34.0%; Pred. No. 1.1e-95;
Matches 334; Conservative 171; Mismatches 361; Indels 117; Gaps 37;

QY 35 CVLINHNKNSVSGKSSYPMVSVESPELGLCARPQSGTGYEAAAEVDVSATTLQVL 94
DB 12 CVLLVLLRGQTGTQSPASGCEPSP-----SIQPAQS-----ELIVEAGDTIRLT 57
QY 95 VDAPGNISCLVFKHSLNCPHFDLQNRGVSVVILKMTQ-TQAGEYLLFIQSEATNYT 153
DB 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIREKAEATHGTGYTCVSGSLRSSI 107
QY 154 ILFTVTSIRN-TLLYTLRRPYPRKMNODALVCISESVPEPIVWVLCDSQGESCKEESA 212
DB 108 YVF---VHDPAPVLVLVGLPLFGK-EDNDALVRCPLTDQD-VSNYSLSIEDCKSLPTDLKF 162
QY 213 V-----VKEEKVVLHLEFGTDIRCCARNELGRECTRLFTID-----LNQTPQTLTLPOL 260
DB 163 VPKAGITIKNVKRAYHRLC---IRCAAQREGKWMRSKFTLKVRAAIKAIPIVVSVPET 219
QY 261 --FLKVGEPILWIRKAVHVNHGFLTW-----ELENKALEEGNYFEMS--TYSTNRTMIR 311
DB 220 SHLLKEGDTFTVICTIKDVTSTVDSMWIKLNPQPOKQVKNRNSWHQGDFFNYERGETILT- 278
QY 312 ILFAFVSSVARNDTGYYTCSSSKHPSQSALVT---IVCKGFIN---ATNNSDEYIDQYE 365
DB 279 -----ISSARVNDSCVFCMYANNFSGSANVTTLKVKVEGFNIFPVKNTT--VFVTGCE 331
QY 366 EFCFSVRKAYPQ-IRCTWTFSRKSPCEQKGLD----NGYSISKFCNHK-----QP 413
DB 332 NVDLVVEFEAYPKPEHQOQWIYMNRT--PTNRCEDYVKNQSNIRYVNELRLRLKLTGTEG 389
QY 414 GEYIFHAENDDAQFTKMTLNIRKPKQVLA--EASASQASCSFDGYPPLPSWTWKCKSDKS 471
```


RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RL pig.";
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223230; CAA1198.1; -.
DR HSSP: P11362; IFKG.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig-like.
DR InterPro: IPR003606; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
KW Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964 964
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;

Query Match 23.9%; Score 1261.5; DB 6; Length 964;
Best Local Similarity 33.4%; Pred. No. 2.2e-95;
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;

QY 47 VKSSSPYMWSESPEDLCALRQSGTGVVEAAVEVDVSASITLQVLVDAPGNISLWV 106
DB 20 VQTGSPQSV--SPEL-----SPSIHPAKS-ELIVSAGDEIRLFCFDPGSVK--WT 67

QY 107 FKSSLLCQPHFDLQNRGVVSMVLKMTQTQAGEYLLFIQSEATNTILFTVSRN--TLL 165
DB 68 FETLG-----QLSENTAEWIEVEAEAMNTGNYCTNEGLSSSIYVF--VRDPEKL 117

QY 166 YTLRRPYFRKMENQDALVCISESPYPIVVEVLCDQSQCKEESPAV-----VKKE 217
DB 118 FLVDPPLYCK-EDNDALVRCPDLPDE-VTNYSLTGCEGKPLPKDLTFVADPKAGITIRNV 175

QY 218 EKVLEHFGTDRC--CARNELGRE-CTRLFTID-----LNQTFQTLTPQ--LFLKVGEPL 268
DB 176 KREYHRL-----CLHCSANQGGKSVLSKFTTLKVRRAIRAVPVAVVASKASYLLREGGEF 229

QY 269 WIRKAVHVNHGFLGWELN-----KALEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDT 325
DB 230 AVNCLIKDVSSVSDSWIRENSTKQAVKNSHWQGFN-----FLQEKLTISSARVND 285

QY 326 GYITCSSKHPQSALVT-----IVGKGFINA-----TNSSEYDEIDQYEFPCFSVRF 373
DB 286 GVEMCYANNTFGSANVTTLLEVVDKGFNIFPMNMTTVFVNDGEDVDL-----IVEY 337

QY 374 KAYPQ-----IRCTWFSRSFCEQGLDNGYSISKFNHKK-----QPGYIIF 418
DB 338 EAYPKPEHRQWYIMNRTATDKWEDPKSE-----NESNIRYSELHLTRLKGTGGGYTF 392

QY 419 HAENDDAQFTKMTFLNIRRRKQVLA--EASASQASCFSDGYPPLPSWTKKCDKSPNCTE 476
DB 393 LVSNDVNSVTNVYVNTKPEILTHDLRMNGMLQCVAAGFPEPTIDWYFCPGTEQRCVS 452

QY 477 EI--TEGVWNRKARKVFGQWSSSTLMNSEAIFKGLVKCCAYNSLGTSCETILNPSGPF 535
DB 453 PVGPVDVQIQNSSVSPGKLVIIHSDISYAFKNGTVECRAYNDVGKS--SAFEN---F 506

QY 536 PFIQD-----NISFYATTGCVLLFTLVLTLLICHYKKQFRIESQLQMV-QVTGS 584

DB 507 AFREQIHAHTLTPLLIGFVIAAGMCIIVMLT-----YKYLQPMYEVQWKVVEEING- 561
QY 585 SDNEYFYDFREYEDLKWEFFPRENLEFFGVLGSGAFGKVMNATAYGISKTGVSQVAVK 644
DB 562 --NNYVVIDPTQLPYDHKWEFFPRNLSFGKTLAGAGFKVVEATAYGLIKSDAAMTVAVK 619
QY 645 MLKEKADSSEREALMSKMTQLGSHENIVNLLGACTLSGPYILIFEYCCYCGDLLNLYR 704
DB 620 MLKPSAHLTEREALMSKLVLSLGNHMINVNLGACTIGGPTLVITVEYCCYCGDLLNLYR 679
QY 705 SKREKPHRTWTE-----IFKEHNFSPYTFQSHPNSSM---PGSREVQIHPDSDQISGL 755
DB 680 RKRDSFCSKQEDHAEALYKNLLHKSKESSCSDSNEMDMKPGVSIV-VPTKADKRRA 738
QY 756 HGNSPHSEDEIYEYENQKRLEEDLNVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAAR 815
DB 739 RIGSY-----IERDVTPAIMEDDEL-ALDLEDLLSFSYQVAKGMAFLASKNCIHRDLAAR 792

QY 816 NVLVTGKVVKICDFGLARDIMSDSNVYVRGNARLPVKNWAPESLPFGIYTIKSDVWSYG 875
DB 793 NILTHGRITKICDFGLARDIKNDSNVVVRGNARLPVKNWAPESIFNCVYTFESDVWSYG 852

QY 876 ILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATETIYIMQSCWAFDSRKRPS 935
DB 853 IFLWELFSLGSSPYPGMPVDVKFYKMIKEGFRMLSPHAPVEMIDIMKTCWDADPLKRPT 912

QY 936 FPNLTSLFGQLADAEAEAMYNQVNDGRVSEC-PH 967
DB 913 FKQIVQLIEKQISESTNHIYSN----LANCSPH 941

RESULT 7
Q8WN23 PRELIMINARY; PRT; 979 AA.

AC Q8WN23;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE C-KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN .[1]
RP SEQUENCE FROM N.A.
RA Zemke D., Yuzbasiyan-Gurkan V.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF448148; AAL40833.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003606; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 4.
DR SMART: SM00410; IG_like; 2.
DR SMART: SM00220; S_TK; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; UNKNOWN_1.
SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DEB8E33D3 CRC64;

Query Match 23.7%; Score 1250.5; DB 6; Length 979;
Best Local Similarity 33.6%; Pred. No. 1.9e-94;
Matches 334; Conservative 161; Mismatches 345; Indels 153; Gaps 39;

```
Qy 47 VGKSSYPMVSPEDLGCALRPQSSGTVEAAAEVDVVSASITLQVLVDAPGNISCLWV 106
Db 22 VQTGSSQSVSPGSLP-SIHPAKS-----ELIVSGDELRLSLCTDGFVK--WT 69
Qy 107 FKHSLLNCQPHFDLQNRGVSMVILKMTETQAGYLLFIQSEATNYTL-----FT 157
Db 70 FETLGLNENTH-----NEWITE-----KAEAG-----HTGNYTCTNRDGLSRSY 110
Qy 158 VSIRN-TLLYTLRRPFYFRKMENODALVCISEVPPIVENVLWCDQSGESKEEPV--- 213
Db 111 VVFRDPAKFLVDLPYLGK-EGNDTLVRCPLTDPE-VTNYSLRGCEGKPLPKDLTFVADP 168
Qy 214 -----VKKEEVLHFLGTDIRC--CARNELGRE-CTRLFTID----LNQTPOTTLPOL- 260
Db 169 KAGITIRNVKREYHFL-----CLHCSADQKGRTVLSKFKFLKVAARAVPVSVSKTS 222
Qy 261 -FLKVGPELWIRCKAVVNHGFLTWELN-----KALEEGNYFEMSTYSTNRTMIRILFA 315
Db 223 SLLKEGEAFVCMCFIKDVSSEFVDSMWIKENSQOTNAQTOSNSHHGDFENFERQEKLI--- 279
Qy 316 FVSSVARNDTGYTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEDEYDQ 363
Db 280 -ISSARNVDSGVFCYANNTFGSANVTTLFVWDKGFNIIPMMSTTTFVNDGENVDL-- 336
Qy 364 YEEFCFSVRFKAYPO-IRCTWTFSRKSPCKQKGL---DNGYSISKFCNKKH------Q 412
Db 337 -----IVEEAYPKPEHQHIYWNRTTDRKWDYPKSDNESNI-RYVSELUHLTLKNE 389
Qy 413 PGEYIFHAENDDAOFTKMTLINIRKPOVLAEASQA---SCFSDGYPLPWTWKCKSDK 470
Db 390 GGTTFQVNSDVNSVTFNVYVNTKPBILTHESLTNGMLQCVAGFPPEPAVDWYFCPGA 449
Qy 471 SPNCTEEL-TEGVNRRANRKFQWQVSSSTLNNSEATKGLVKCCAYNSLGTSCETILL 529
Db 450 EQRCVPIGPMQVQMSLSPSGKLVVQSSIDYSAFKHNGTEVCRAYNNVGRS--SAFF 507
Qy 530 NSPGPPFPFIQDN-----ISFYATIGVCLLFIIVVLTLLICHKVKKQPRYEQ 575
Db 508 N----FAFKGNSKEQIHPHTLFTPLLLGTVIAGMNCIIVMILT-----YKYLQKPMYEVQ 559
Qy 576 LQWV-QVTGSSDNSEYFYVDREYEDLKWEPFRENLEFGKVLGSGAFGKVMNATAYGISK 634
Db 560 WKVVEEING---NNVYVTDPTQLPDYDKHWEPPRNLSPGKTLGAGAFKVVVEATAYGLIK 616
Qy 635 TGVSIQAVKMLKEKADSSREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIFEYC 694
Db 617 SDAAMTVAVKMLKPSAHLTEREALMSLKVLSYLGNNHNMIVNLLGACTVGGPTLVITEYC 676
Qy 695 CYGDLNLYLRKSKREKFKHRTWTE-----IPKEHNFSEFYPTFQSHPNSSM--PGSREYQI 745
Db 677 CYGDLNLYLRKSKREKFKHRTWTE-----IPKEHNFSEFYPTFQSHPNSSM--PGSREYQI 735
Qy 746 HPDSQISGLHNSGFHSEDETEYENQKRLEBEDNLVLTFFEDLLCFAYQVAKGMEFLBFX 805
Db 736 PTRADKRRSARIGSY-----IERDVTPTAIMEDEL-ALDLEDLLSFSQVAKGMAFLASK 789
Qy 806 SCVHRDLAARNVLTGKVKYICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGYI 865
Db 790 NCVRDLAARNVLTGKVKYICDFGLARDIKDSDSNVYVRGNARLPVKWMAPESLFEGYI 849
Qy 866 TIKSDVWSYGLILLWEIFSLGVNPPGIPVDANFYKLLQNGEKMDOFPYATEIYIIMQSC 925
Db 850 TFDSDVWSYGLIFLWELSLGSSPPGMPVDSKFKYMKIEGFRMLSPHAPAEYDIMKTC 909
Qy 926 WAFDSRKRPSFPNLTSLGCOLADAEAMYNQV 958
Db 910 WDADPLKRPTEFKQIVLJEKQISDTHIYSNL 942
RESULT 8
Q99662
ID Q99662 PRELIMINARY; PRT; 972 AA.
AC Q99662;
```

```
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE KIT protein.
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92228497; PubMed=1373482;
RA Andre C., Martin E., Cornu F., Hu W.X., Wang X.P., Galibert F.;
RT "Genomic organization of the human c-kit gene: evolution of the
RT receptor tyrosine kinase subclass III."
RL Oncogene 7:685-691(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachauve P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes."
RL Genomics 39:216-226(1997).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; U63834; AAC50969.1; -.
DR HSSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferrase; Transmembrane; Tyrosine-protein kinase.
SQ SEQUENCE 972 AA; 109450 MW; D59DFE9AF761FDA CRC64;
Query Match 23.6%; Score 1247; DB 4; Length 972;
Best Local Similarity 32.8%; Pred. No. 3.6e-94;
Matches 323; Conservative 166; Mismatches 353; Indels 144; Gaps 34;
Qy 47 VGKSSYPMVSPEDLGCALRPQSSGTVEAAAEVDVVSASITLQVLVDAPGNISCLWV 106
Db 20 VQTGSSQSVSPG-EPSPPSIHPGKSDLI-----VRVGEIRLLCTDGFVK--WT 67
Qy 107 FKHSLLNCQPHFDLQNRGVSMVILKMTETQAGYLLFIQSEATN---YTLFTVSIKNT 163
Db 68 FE-----ILDETNEKNQEWIT-EKAEATNTGKYTCTNKHGLSNS 106
Qy 164 L-----LYTLRRPFYFRKMENODALVCISEVPPIVENVLWCDQSGESKEE----- 209
Db 107 IYVFRDPAKFLVDLPYLGK-EDNDTLVRCPLTDPE-VTNYSLRGCEGKPLPKDLRFIP 164
Qy 210 ---SPAVVKKEEVLHFLGTDIRCARNELGRECTRLFTIDL---NOTPOTTLPO--L 260
Db 165 DPKAGIMKSKVRAVHRLC---LHCSVDQEGKSVLSEKFLKVRPAFRAPVSVSKASY 221
Qy 261 FLKVGPELWIRCKAVVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFV 317
Db 222 LLREGEETVCTIKDVSSSVYVSTWKRNSQTKLQEKYNSHHGDFNYERQAT---LTI 277
Qy 318 SSVARNDTGYTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEDEYDQYE 365
Db 318 SSVARNDTGYTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEDEYDQYE 365
```

```

Db 278 SSARVNDGSGVEMCYANNFTGSGANVTTLTLEVVDKGFINFPMINTTFVNDGENVDL---- 333
Qy 366 EFCFSVRFKAYPO-IRCTWTFSRKSP--PCEQKGLDNGYISIFCNHKKH-----QPGE 415
Db 334 ----IVEYEAFPRKPEHQOQWYIMRFTDKWEDYPKSENEINIRYVSELHLTRLKLGTEGGT 389
Qy 416 YIFHAENDDAQTKMFTLNIRRRKPOVLA--EASASQASCFSDGVPPLPSWTWKCKSDKSPN 473
Db 390 YTLVLVNSDVNAIAFNYYVNTKPEILTYDRLVNGMLQCVAAAGEPEPTIDWYFCPGTEQR 449
Qy 474 CTEEITE-GVWNRKANRKYFGOWSSSTLNMSEALKGLFKCCAYNSLGTSCETILLNSP 532
Db 450 CSASVLPVDVQTLNSSGPPGKVLQVQSSIDSSAFKHNGTVECRAYNDVGKT--SAYEN-- 505
Qy 533 GPFPFTQD-----NISFVATICVCLLFIVVLTLCHYKKKFRYESQLQWV-QV 581
Db 506 --FAFKEQIHPHTLFTPLLLGFGVIVAGMCIIVMLT-----YKYLQKPMVEQWKKVVEEI 559
Qy 582 TGSDDNEYFYVDREYEDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSTQV 641
Db 560 NG---NNYVYIDPTQLPYDHKWEFPRNRLSEFKTLGAGAGKVVVEATAYGLIKSDAAMTV 616
Qy 642 AVKMLKEKADSSEREALMSELKMTOLGSHENIYNLLGACTLSGPIYLIPIFYCCYGDLLN 701
Db 617 AVKMLKPSAHLTEREALMSELKVLVSLGNHNMNIYNLLGACTIGGPTLVITEYCCYGDLLN 676
Qy 702 YLRSKREKFRHTWTE-----LKEHNFSYPTQSHPNSSM----PGSREVQIHPDSQI 752
Db 677 FLRRKRDSFICSKQEDHAEALYKNLLHRSKSSCSDSTNEYMDMKPGVSYV-VPTKADKR 735
Qy 753 SGLHNSFHSDELEYENQRLKEEEDLNVLTFEDLLCFAYOVAKGMFELEFKSCVHRDL 812
Db 736 RSVRIGSY-----IERDVTPAIMEDDEL-ALDLEDLLSFYQVAKGMAFLAKNCIHRDL 789
Qy 813 AARNVLVTHGKVKVLCDFGLARDTMSDSNVYVRGNARLPVKWMAPESLFEGIYTIKSDVW 872
Db 790 AARNILTHGRYIKCDFGLARDKNSNVYKGNARLPVKWMAPESLFENCYVFESDVW 849
Qy 873 SYGILLWEIFSLGPNYPGIPVDANFYKLIQNGKMPQDPFYATEEYIIMOSWAFDSRK 932
Db 850 SYGIFLWELFSLGSPYPGMPVDSKFKYMKIEGFRMLSPHAPAEYMDIMKTCWDADPLK 909
Qy 933 RPSFPNLTSLGCOLADAEAEAMQV 958
Db 910 RPTFKQIVQLIEKQI1SESTNHIYSNL 935

RESULT 9
Q9XS93
ID Q9XS93 PRELIMINARY; PRT: 978 AA.
AC Q9XS93;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE K.T.
GN C-KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RA London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,
RA Gelsler E.N.;
RT "Spontaneous canine mast cell tumors express tandem duplications in
the proto-oncogene c-kit.";
RL Exp. Hematol. 0:0-0(1999).
CC C-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF090303; AAD28369.1; -.
DR HSSP; P11362; 1FGK.

```

```

DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RtkinaseI1.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 2.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_ILI; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
SQ SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

```

Query Match 23.6%; Score 1244; DB 6; Length 978;

Best Local Similarity 33.2%; Pred. No. 6.4e-94;

Matches 330; Conservative 161; Mismatches 346; Indels 158; Gaps 38;

```

Qy 47 VGKSSYPMVSESPEDLGCALRPQSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLMW 106
Db 22 VQTSSQPSVSPGPELP-SIHPAKS-----ELIVSVGDELRLSCTDPGFVK--WT 69
Qy 107 FKH-SSLNCQPHDLQNRGVVSMVLKMTQAGLYLLFIOSEATNYTIL-----FT 157
Db 70 FETLGQLNENTH---NEWITE-----KAEAG-----HTGNYTCTNRDGLRSIY 110
Qy 158 VSIRN-TLLYTLRPRYFRKMENQDALVCISEVPPIVENVLVCDSQGESCREESAV--- 213
Db 111 VFVRDPAKLFVLVDLPYLGK-EGNDTLVRCPLTDP-VTNYSLRGCEKPLPKDLTFVADP 168
Qy 214 ----VKKEEVKLVHELFGTDIRC--CARNELGRE-CTRLFTID----LNQTPQTLPOL- 260
Db 169 KAGTIRNVKREYHRL-----CLHCSADQKGRTVLSKFKTLKVRAAIRAVPVSVSKTS 222
Qy 261 -FLKVGPELWIRCKAVHNVHGFGLTWELN---KALEEGNYFEMSTYSTNRMTIRILFAF 316
Db 223 SLLKEGEAFVCMFTKDYVSFVDSMWIKENSQTNAQTQSNHSHHGDENFEROEKLI---- 278
Qy 317 VSSVARNDTGYITCSSKHPQSQALVT---IVGKGFINA-----TNSDEYEDIDY 364
Db 279 ISSARVNDGSGVEMCYANNFTGSGANVTTLTLEVVDKGFINFPMINTTFVNDGENVDL--- 335
Qy 365 EEFCSVREKAYPO-IRCTWTFSRKSPCEOKGL---DNGYSLSKFCNKKH-----Qp 413
Db 336 ----IVEYEAYPRPEHQOQWYIMRFTDKWEDYPKSDNESNI-RYVSELHLTRLKNGEG 389
Qy 414 GEYIFHAENDDAQTKMFTLNIRRRKPOVLAESAQA--SCFSQGYPLPSWTWKCKSDKS 471
Db 390 GTYTFQVNSDVNSVSTENVYVNTKPEILTHESLTNGMLQCVVAGPEPAVDWYFCPGA 449
Qy 472 PNCTEEI-TEGVWNRKANRKYFGOWSSSTLNMSEALKGLFKCCAYNSLGTSCETILLN 530
Db 450 QRCSPYIGPMDVQMQNSLSLSPGKLVQSSIDYSFAKHNGTVECRAYNNVGRS---SAPFN 507
Qy 531 SPGPFPFTQD-----ISFVATICVCLLFIVVLTLCHYKKKFRYESQL 576
Db 508 ----FAFKGNSKEQIHPHTLFTPLLLGFGVIVAGMCIIVMLT-----YKYLQKPMVEQW 559
Qy 577 QWV-QVTSDDNEYFYVDREYEDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKT 635
Db 560 KVVEEING---NNYVYIDPTQLPYDHKWEFPRNRLSEFKTLGAGAGKVVVEATAYGLIKS 616
Qy 636 GVSIQAVKMLKEKADSSEREALMSELKMTOLGSHENIYNLLGACTLSGPIYLIPIFYCC 695
Db 617 DAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNMNIYNLLGACTIGGPTLVITEYCC 676
Qy 696 YGDLNLNLSRKREKFRHTWTEIFKEH-NFSFYPTQSHPNSSMPGSRREVQIHPDSDQISG 754

```

```
Db 777 YGDLNLFRRKRDSTFCRQKQ---DHGEVALYKNLLHSSKSSDS-----TNEYMDM 726
QY 755 LHGNSPHSDEIEYENQKRLKE-----EEDLNVLTFFEDLLCFAYQVAKGMEFLE 803
Db 727 KGDSTVVTPTKADKRKRRIIGSVIERDVPAINMEDDELALDLELUSFSQVAKGNAPLA 786
QY 804 FKSCVHRDLAARNVLVTHGKVKYKICDFGLARDIMSDSNYYVRGNARLPVKWMAPESLFEG 863
Db 787 SKNCIHRDLAARNILLTHGRITKICDFGLARDIKNSYVVKGNARLPVKWMAPEISFNC 846
QY 864 IYTIKSDVMSYGIILLWEISLGVNVPVPGIPVDANFYKLQNGFKMDQPPVATEEIIIMQ 923
Db 847 VYTFESDVMSYGIIFLEFSLGSPVPGVPVDSKFKYKMKIEGFRMLSPHAPAEYDIMG 906
QY 924 SCWAFDRKRPSFPNLTSLFGCOLADAEERAMQNV 958
Db 907 TCWADAPLKRPTSKQIVQLEIKQISDSTNHIYSNL 941

RESULT 10
Q98SUI PRELIMINARY; PRT; 977 AA.
AC Q98SUI;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Receptor tyrosine kinase Fms.
OS Danio nigrofasciatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=144739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Patrachy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
diversification in danio."
RL Dev. Genes Evol. 211:319-328(2001).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
CC EMBL: AF324481; AAK15303.1; -.
DR HSP; P11362; 1FGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RtkinaseII.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_4.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; Ig; 4.
DR SMART; SM00408; IgC2; 1.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferrase; Transmembrane.
FT VARIANT 60 60 V -> L.
FT VARIANT 103 103 A -> S.
FT VARIANT 742 742 P -> S.
FT VARIANT 899 899 E -> D.
SQ SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;
```

```
Query Match 23.3%; Score 1230.5; DB 13; Length 977;
Best Local Similarity 32.8%; Pred. No. 8.4e-93;
Matches 326; Conservative 174; Mismatches 367; Indels 127; Gaps 34;

QY 71 SSGTVYEAAVEVDVSASITLQVLVDAPGNISCL-WYFKHSSLNCQPHDLQNGVYWSV 129
Db 27 NSGAL---AGTDVILDSGSLQLVCEGDPVTFPPRAKH-----KRYISKEVGKIRSF 78
QY 130 ILKMTETQAGEY-LLFQSEATNYTILFTYSIRNT-LLY-----TLRPPYERKVENQDAL 182
Db 79 VERATVDFTGYKCVIYINGSDSNLASSVHVFRDSRVLVFVSPSSSLR--YVRK-EGEDLL 135
QY 183 V-CISESVPEP-IVEWVLCDSQGESKEESPVVVKKEKVL-----HELFGTDIRCCARNE 236
Db 136 LPCL---LTPD DATFTFRMDNGSAAPYGMNATDPKGVLI RNVHFGFNADYIC SARIG 192
QY 237 LGRECTRLFTIDLNQTPQTTLPLQFLK-----VGEPLWIRKAVHVNHGFGITWELN 289
Db 193 GAEKVS KIFSI NVIQR LRP- P YVY LKRNE YV KLVGR LQISCTTNNP N FYI NV TWTHSS 251
QY 290 KALEEGNVFEMSTYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSQALVT---IVG 346
Db 252 KRLPKPE--EKSTMEDRLAIESILT-ISSVQLSDTGNITCTGQNEAGANSTQLLVVD 308
QY 347 KGFI-----NATNSSEYEDIDQEEFCFSVRFKAYPOIRC-TW---TESRKSFPCEQ 394
Db 309 EPYIRLSPKLSKLTGRGLSIEVSEGEDVDLGVLEIAYPLTSHTWETPTSHNASLP--- 365
QY 395 KGLDNGYSISKFCNHK-----HQGEYIFHAENDDAQTKMFTLIRKPKQ 440
Db 366 -----ENRFYNHNDRYEALLFLKRLNFEIEIGQYTLNVKNMSKASITFDIKWYTKPV 417
QY 441 VLAE-ASASQASCFSDGYPLPSTWTKKCDKSPNCTEITE-----GVNKRANKRVFGOW 495
Db 418 ARYKWNVTLLSCRSYGPAPSLWYQCTGIRTCPTENTDLOPIQTQTVFQKSEGA 477
QY 496 VSSSTLNMSEAIGFLVKCCAYSLGTSCTETILLNSPPFPFIODNTSFYATIG--VCLL 553
Db 478 GVSEVLTVGN-RRMTVVCVAFNVLVGOGSDTFSD-----VSDQFTSAMCSTVAMV 529
QY 554 FIVVLTLLICHYKQPRYESQLQMVQVQVTSNDNEYFYDFREYEDLKNFPPRENLEFG 613
Db 530 VLGLLLIFMIYKYKQPRYEIRKIIIEAT--NGNNYTFIDPTQLPYNEKEWFPDRDKLIG 587
QY 614 KVLGSAGFGKVMNATAYIGISKTGVSIQVAVKMLKEKADSSERREALMSELKMMTOLGSHEN 673
Db 588 KTLGAGAFGKWEATAYGLGKEDNITRVAVKMLKASAPDEREALMSELKILSHLGQHN 647
QY 674 IVNLGACTLSGPITYLIFEYCCYGDLLNLYLSRKREKPHRTWTEIFKEHNFSEFYTFQSH 733
Db 648 IVNLGACTHGGPVLVITEYCCHGDDLNFURSKAENF-----LNFVMTI---P 692
QY 734 NSSMP-----GSREVOIHPDSDQISGLHGNPSHSEDEIEYENQKRLKE-----EEDLNV 782
Db 693 NFPEPVTDYKNVSTERMEFVRSDSGISCTSDHYLDMPVTSRPTNSALDPSDDCQESWP 752
QY 783 LTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKYKICDFGLARDIMSDSN 842
Db 753 LDMDDLLRFSQVAGQLDFLAANKNCIHRDVAARNVLLTNSRVAKICDFGLARDIMSDSN 812
QY 843 VVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGIILLWEISLGVNVPVPGIPVDANFYKLI 902
Db 813 VVRGNARLPVKWMAPEISIFECVYTVQSDVMSYGIIMLEIFSLGKSPYPNILVDSKFKMI 872
QY 903 QNGFKMDQPPVATEEIIYIMQSWAFDRKRPSFPNLTSLFGCOLADAEERAM---XQNV 958
Db 873 KCGYQMSRPDPFAPEMYTINKMCWNLEAERPTFSKISQMIQRMGLSETSEQQDTQYRKNI 932
QY 959 DGRVSECPHTYQNRPRPSREMDLGLLSPOAQVED 992
Db 933 PSE-AEVEQPLESCDPVKHDESEFETSCEPEED 965
```

```

RESULT 11
ID 077589 PRELIMINARY: PRT: 945 AA.
AC 077589;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Tyrosine kinase receptor homolog (Fragment).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Lear T.L., Adams M.H., McDowell K.J., Ponce de Leon F.A.,
RA Sullivan N.D., Coogler L., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, KIT, MX1 and PGR
RT in the horse, Equus caballus.";
RL Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RT "An Equine sequence homologous to the c-KIT proto-oncogene mapped to
RT chromosome 3q13-21.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF055037; AAC23611.1; -.
DR HSSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Glycoprotein; Immunoglobulin domain; Kinase; Phosphorylation;
KW Receptor; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 945
SQ SEQUENCE 945 AA; 106143 MW; 7DB424A89DF5C340 CRC64;

Query Match 23.3%; Score 1229; DB 6; Length 945;
Best Local Similarity 32.5%; Pred. No. 1 le-92;
Matches 323; Conservative 166; Mismatches 368; Indels 136; Gaps 31;

QY 47 V GKSSYPMVSEPDGALRQSSGTVEAAAEVDVSAITLQVLVDAPGNISCLAW 106
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20 VQTGSSQPSV--SPGEL-----SPSIHPAKS-ELIVSGVEIRLLCADPGFVK--WT 67

QY 107 FK-HSSLNCPHFEDLQNRGVSVVILKMTETQAGEYLLFIQSEATNTYTLFTVSIINTLL 165
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 68 FETWGLSENTHKE-----WVTEKAEATNTGVTCTNEGSLSIYVF---VRDPAK 116

QY 166 YTLRRPYFRKMNODALVCISESPYPIVEWVLCDQSGESCKEESPAV-----VKKE 217
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 LFLFDSLPGKSSDYLVRCLPTDPE-VTNYSLMACEGKSLPKDLTFVADPKAGITIRNV 175

QY 218 EKVLFHFGTDIRCCARNELGRECTRLFTID----LNQTPQTTLQ--LFLKVGGEPLWIR 271
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 KREYHRLC---LRCSADKQGSVLSNFTLKVRAAIVRAVPVSVKASYLLREGEEFSVT 232

```

```

QY 272 CKAVHVNHGFLTWELNKALEEGNYFEMSTYSTNRTMIRLFAFVSVSVARNDTGYTCS 331
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 SSKHPSQSALVT---IVKGFINA-----TNSSEDEIYOYEFCEFSVRFKAYPQ- 378
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 ANNTFGSANVTTLTLEVLDKGCIINVFPMMNTTVFVNDGENVDL-----IVEYESYKPK 343
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 379 -----IRCTWTFSRKSPFCEQKGLDNGYSISKFCNKH-----QPGEYIHAENDD 424
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 EHQWIYMNRTSTDKWEDYPKSE-----NESNIRYVSELHLRLKLTGEGGTYYTLVNSND 398
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 425 AQFTKMFLLNIRRKPOVLA--EASASQASCFSDGYPSPMTWKCKSDKSPNCTBEI--TEG 481
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 399 VDSSTVFNVYNTKPEILTRDLRMNGMLQCVAAAGFPPTIDWYFCPTGEQRCSPVGPVD 458
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 VNRKANRKFQGWVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPFFIQDN 541
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 VKIQNSSVSPFGLVQSSIDYSAFKHNGTVCECRAYNDVGKSAFF-----N 505
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 ISFYATIGVCLLF-----IWLTLCHIKYKQFRYESQLOWV-QVTGSSDN 587
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 FAFKEQIHPHTLFTPLLIGSVAAAGMCMVIMVLTLYKLOKPMYEVQMKVVEEING---N 562
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 588 EYFYVDREFEYDLKWEFPRENLEFGKVGAFKVMNATAYGISKTGVSIOVAVKMLK 647
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 NYVIIDTQLPYDHKWEFPNRNLSFGTLGAGAFKVVEATAYGLIKSDAAMTVAVKMLK 622
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 648 EKADSSEREALMSKMTQLGSHENIVNLGACTLSGPIYLIFEYCYGDLNLYLSKR 707
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 PSAHLTEREALMSKLVLSYLGHNHNVNLLGACTVGTPLVITEYCYGDLNLYLSKR 682
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 708 EKPHRTWTE-----IPKEHNFSEYPTFQSHNSM---PGSREVQIHPSDQISGLHGN 758
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 DSFTCSKOEDHAEALYKLLHLSKSSCNDSTNEYMDMKPGVSYV-VPTKTKRRAARIG 741
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 759 SFHSEDEIYENOKRLEEDLNLTFTEDLLCFAYQVAKGMFELEKSCVHRDLAARNVL 818
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 742 SY-----IERDVAPSIMEDDEL-ALDLEDLLSF SIQVAKGNFLASKNCIHRDLAARNIL 795
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 819 VTHGKVVKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILL 878
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 LTHGRITKICDFGLARDIKNSNVVKGNVRLPVKWMAPESIFNCVYTFESDVWSYGIFL 855
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 879 WEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPPYATEEYIIMOSWAFSPRRKRPSPN 938
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 856 WELFSLGSSPYPGMPVDSKFKYKIEGFRMLSPHAPAEYDMKTCWDADPLKRPTFKQ 915
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 939 LTSFLGCOLADAEAEAMYQNDGRVSECPHYON 971
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 916 IVOLIEKQISDSTNQIYSN-----LANCSPROEN 944

RESULT 12
Q98SU3 PRELIMINARY: PRT: 974 AA.
AC Q98SU3;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase Fms.
GN FMS.
OS Danio dangila.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=127599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT diversification in Danio.";

```

[illegible]

470 EGSEFGSGVGVESLVTGPN--RMTVVVCAFNVLVQCQSDTFSMD-----VSDQIFTSAM 521

548 IG--VCLLFIVTLTICHKKYQFRYESQLMVQVVTGSSDNEYFVDFREYEDYDLKWEF 605

552 CGSTVAMVVLGGLLIPMIYKKYKQKPRYEIRWKIIEAT--NGNNYTFIDPTQLPYNEKWEF 579

606 PRENLFGVLGSGAFGKYNATAYGISKTGVSIQVAVKMLKEKADSSREALMSCLKMM 665

580 PRDLKLGKTLGAGAFGKVVYEATAYGLGKEDNITRVAVKMLKASAHPDREALMSCLKIL 639

666 TOLGSHENIVNLGLACTSGPIVLIFFEYCCYCGDLLANLRSKREKPHRTWTEIFKEHNFSEF 725

640 SHLGCHKNIVNLGLACTHGGFVLITVEYCCGDLNLFNLSKKAENF-----LNLF 687

726 YPTFQSHPNSSMP-----GSREVOIHPDSQISGLHGNFSHSEDEIEYENQKRLEE----E 777

688 VMTIPDFPEPMTDYKNISTERMEFVRSDSGTSSAC--SDHYLDMRPVTSRPTLDSSEQC 744

778 EDLNVLPEDLLCFAYOVAGMGELFEKSCVHRDLAARNVLVTHGKVKVICDFGLARDIM 837

745 EDSVPLMDMLLRFSSQVAGLDFLAAKNCIHRDVAARNVLLTNSVAKICDFGLARDIM 804

838 SDSNVVYRGNARLPVKWMAPESLFEGITYTIKSDWSVYGLLWEIFSLGVNYPGIPVDAN 897

805 NDSNVVYGNARLPVKWMAPESEFECYTVQSDWSVGTGMLWEIFSLGSKSPYPNILVDSK 864

898 FYKLIQNGFKMDQPFYATEIYTIIMOSCAFDKRPSPNLTSLFGLCOLADAEAM--- 954

865 FYKMKCGYQMSRDFAPPENYTIKMCWNLDAERPTFSKISQMTQRMLETPEQOETQ 924

955 -YQNV-----DGRVSECPHYQNRRPF-----SREMDLGLLSP 986

925 EYKNIPSEAEAEATQLESCDPAKHEDSEFETSCDQEEEDQPLMKP 968

RESULT 13

Q91909 ID Q91909 PRELIMINARY; PRT; 954 AA.

AC Q91909;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE C-kit-related kinase 1 (Xkrk1) precursor.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

ON NCBI_TaxID=8355;

OX [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95344996; PubMed=7619732;

RA Baker C.V., Sharpe C.R., Torpey N.P., Heasman J., Wylie C.C.;

RT "A xenopus c-kit-related receptor tyrosine kinase expressed in

RT migrating stem cells of the lateral line system.";

RL Subcell. Dev. 50:217-228(1995).

CC -!- SUBCELLULAR LOCATION: TYPE I

CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE

CC PROTEIN KINASES.

DR EMBL; Z48770; CAA86868.1; -.

DR HSP; P11362; IFGK.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003600; Ig_like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR001824; RTKinaseIII.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00047; ig; 2.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Euk_pkinase; 2.

DR SMART; SM00410; IG_like; 2.

DR SMART; SM00319; TyrcK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

Db 254 LPKAE--EKSTMEGRDIAIESILT-IPSVLSHTGNITCTQNEAGANSSYTLQVLVVEEP 310
QY F1-----NATNSEDYEDIDQVEECFVSFRFAYQIIRC-----TWTFSRKSPCEQKG 396
Db 311 YIRLSPKLSKLTGRGLSIEVSEGDVDLGLVLIAYPLTSHKWTPTSHNASLP----- 365
QY 397 LDNGYSTSKFCNHK-----HQPGEYIFHAENDDAQFTKMTNIRKKPOVL 442
Db 366 -----ENREFNDRYEAALLLRLNFEETIGQYTLVKNKMSASITFDIKMTKPVAR 419
QY 443 AE-ASASQACSFSDGYPLPWTWKCKSDKSPNCTEETE-----GVWNRKANRKFQGVWS 497
Db 420 VKWENVTLSCRSYGPAPSILLTQCIGIRTCPENTDLOPIQTQVVEFKESFGVGV 479
QY 498 SSTLNMSAELKGFVLKVCAYNSLGTSETILLNSPGFPFPFTQDNISFYATIG--VCLLFI 555
Db 480 ESVLTGCPN-RMTVVCVAENLVGSGDFTSME-----VSDQIFTSAMCGSIVAMVVL 531
QY 556 VLTLLICHYKKOFRYESOLQVQVWGSSDNEYFYDFREYEDLKWEPFRENLEFGKV 615
Db 532 GLLIFMIYKOKPRYIEIRWIKIIEAT--NGNNYTFIDPTOLPYNEKWEPRDKLKGKT 589
QY 616 LGSAGFKVMNATAYGISKTCVSTQVAVKMLKERADSEREALMSELKMTQOLGSHENIV 675
Db 590 LGAGAFKVVYEATAYGLGKEDNITRVAVKMLKASHPDEREALMSELKILSHLGQHNIV 649
QY 676 NLLGACFLSGPIYLFYCCYVGLDNLNLSRKREK-----HRT 713
Db 650 NLLGACHTGGPVLVITYCCCHGDLNLSKAENFLNFWMTIPNPEPMTDYKNVSTERM 709
QY 714 W-----TEIFKEHNFSFYPTFQSHNPSMPGSGREVOIHPDSQISGLHGNFSHDEI 766
Db 710 FVRSDSGISSTCSHYLDMRPVTSRPTNSALDSSSEQ----- 747
QY 767 EYENQKLEBEDLNVLTFFEDLLCFATQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKV 826
Db 748 -----EDSWPLDMDDLRFSSQVAGLDFLAKNKIHDRVAARNVLLTNSRVAK 796
QY 827 ICDPLGARDIMSDSNVYVRGNARLPVKKMAPESLFEGLYTIKSDVWSYGILLWEIFSLGV 886
Db 797 ICDPLGARDIMSDSNVYVRGNARLPVKKMAPESIFECVYTVQSDVWSYGIMLWEIFSLGK 856
QY 887 NPYGPIVDANFYKLIONGFKMDQFPYATEEIIYIMOSCAWFDGRKRPSFPNLTSLFLGQ 946
Db 857 SPYPNILDVSKFYKMKCGYQMSRPDFAPEPMYIMKMCNLDAAERPTFSKISQMIQRM 916
QY 947 LADAEAM-----YQNV-----DGRVSECPTYQNRPP-----SREMDLGLLSP 986
Db 917 LGSETSEQDQYKNIPTAEAEAOQLESQDPVKHDESEFETSCDQEEEDQPLMKP 971

RESULT 15

Q9TDT7 PRELIMINARY; PRT; 948 AA.
ID Q9TDT7
AC Q9TDT7
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Stem cell factor receptor (fragment).
GN C-KIT.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichsurus.
OX NCBI_TaxID=9337;
RN [1]
SEQUENCE FROM N.A.
RA Lawrence S.B., Greenwood P., Tisdall D.J., McNatty K.P., Fidler A.E.;
RP "Partial cDNA sequence of the c-kit homologue of brushtail possum
RT (Trichosurus vulpecula).";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131209; AAF22141.1;
DR HSSP; P11362; 1FGK.

DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Immunoglobulin domain; Receptor; Tyrosine-protein kinase.
FT NON_TER 948
SQ SEQUENCE 948 AA; 106323 MW; 708E0258FFC4D07D CRC64;
Query Match 23.2%; Score 1221; DB 6; Length 948;
Best Local Similarity 33.2%; Pred. No. 4.9e-92;
Matches 319; Conservative 160; Mismatches 363; Indels 118; Gaps 32;
QY 67 LRPOSSGTVYEAADVEVDVSASITQLVLVDAPGNISCL-----WVFKHSSLNCQPHFDL 120
Db 24 LQSTSPANYSSPSI---TPARSQLTVMFGSEIKLSCSDLHFVNMTFFENALHLE----- 75
QY 121 QNRGVYVMILKMTETQAGEYLLFIQSEATNYTILFTVTSIRNTLLYLRPYFRKMNQD 180
Db 76 SSRTTETWLTNAKAQDGRYCTCTNKMGLSSSIYVVKDPK--ILFLDLRLVGN-EGSD 132
QY 181 ALVCISESVPEPIVW--LCD-----SQGESKEESPAYVVKKEKVLHEL 224
Db 133 ALVNC-PVTDPE--VTNFTLRLCDGKPLKDLTLIPDIQGITKN-----VKRSHNKIC--- 184
QY 225 FGTDIRCCARNELGRECTRLFTIDLNQTPQTTLPOL-----FLKVGPEPLWIRCKAVHV 277
Db 185 ----PQSAAYQGGUQLSDRMFLKVRPAQ--SVPEVSLQTNLYLREGTFFQATCMKDV 239
QY 278 NHGFLTWELNKALEEGNFEFEMSTYSTRNTRMIRILFAFVSSVARNDGYTSCSSK--- 334
Db 240 ASSVYVMWKNNRI---STHTQSRHSGDYAYERQDILTISPVRVNDSGVTCFANNTG 296
QY 335 HPSGALVTYVKGGINA--TNSSDEYEDIDQVEECFVSFRFAYQ-----TRCTWTFSR 387
Db 297 PANVTATLVKVEKGINFPPKMTTIFINDGENIDLWVEYEAFFKPEHLQWIMNGTVD 356
QY 388 KSFPCQKGLDNGYSISKFCNHK-----OPGEYIFHAENDDAQFTKMTNIRKKPQ 440
Db 357 KWDOYTRKPESE---STIRYISELHLNRKLGTEGGPYTFSVNSDVSTSVTFKYVVKRPE 413
QY 441 VLAESAASQA--SCFSDGYPLPSWTWKCKSDKSPNCTEETEGVWNRKANRV---FGQW 495
Db 414 ILTSDRLNGLLQCVAGFPPEPTIDWYFCPGTEQRCSTSLP--MDVKTNSMLPPFGKI 471
QY 496 VSSSTLNMSAELKGPLVKCCAYNSLGTSETILLNSPGFPF--IQDNISFYATIGVCLL- 553
Db 472 VVESTIDSSAFRYNGTVECKASNDVG-----KSSAFFNFAIKEQITSHLTFTPLLG 523
QY 554 FIV-----VLTLLICHYKKOFRYESOLQV--QVTGSSDNEYFYDFREYEDLKWEEP 606
Db 524 FVAAGLMCVIILTYKIFQKPMYEVQMKVVEEING---NNYVYIDPTQLDYDHKWEPP 580
QY 607 RENLEFFGVLSGAFGKVMNATAYGISKTGVSIOVAVKMLKERADSEREALMSELKMMT 666
Db 591 RNRLSFGKTLGAGAGKVVYEATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSELKVL 640
QY 667 QLGSHENTVNLGACTLSGPIYLFYCCYVGLDNLNLSRKREKFRHTWTFEIKH-----N 722
Db 641 YLGNHMNIANLIGACTIGGPTLVITEYCCYVGLDNLNLSRKROSF-----ICKSHEDHAE 694
QY 723 FSPYPTFQSHNPSMPGSGRE--VOIHPDSQISGLHGNFSHSE---DEIYEYENQKLEEE 778
Db 695 AALYKNLLQSKESCDGANEYMDMKPGSVYVPTKAEKRRRSARVGSYIERDVTYTAIMEOD 754

